ÁLGEBRA LINEAL

CLAUDIO DE J. PITA RUIZ
Escuela de Ingeniería
Universidad Panamericana

Revisión técnica
Fernando Vera Badillo
Ing. Civil, Universidad La Salle
Maestro en Ciencias, UNAM
Profesor Investigador
Universidad La Salle

McGRAW-HILL

MÉXICO • BUENOS AIRES • CARACAS • GUATEMALA • LISBOA • MADRID • NUEVA YORK
PANAMÁ • SAN JUAN • SANTAFÉ DE BOGOTÁ • SANTIAGO • SÃO PAULO
AUCKLAND • HAMBURGO • LONDRES • MILÁN • MONTREAL • NUEVA DELHI • PARÍS
SAN FRANCISCO • SINGAPUR • ST. LOUIS • SIDNEY • TOKIO • TORONTO

ÁLGEBRA
LINEAL

PRÓLOGO

Not being familiar with the concepts of Linear Algebra such as linearity, vector, linear space, matrix, etc. nowadays amounts almost to being illiterate in the natural sciences and perhaps in the social sciences as well.

Lars Gårding

Este libro contiene un curso introductorio de álgebra lineal. Hace algunos años esta disciplina matemática era estudiada solamente por personas que se estaban formando dentro de alguna carrera de "ciencia pura", como matemáticas o física. Es por eso que actualmente existen en el mercado algunos textos escritos precisamente para este público. Estos textos son, sin duda alguna, excelentes libros sobre el tema, pero tienen la característica de ser muy rigurosos o muy ambiciosos en lo que respecta a la generalidad de la teoría tratada.

Más recientemente, el álgebra lineal se ha convertido en una parte de la matemática que resulta ser *indispensable* para cualquier persona que estudia una carrera científica o técnica, como lo es cualquier rama de la ingeniería. Es entonces ahora, que han comenzado a aparecer algunos libros sobre este tema con características distintas de aquéllas que mencionaba anteriormente. La intención de esta nueva generación de textos de álgebra lineal es introducir al estudiante a esta parte de la matemática de un modo "didáctico", partiendo de situaciones concretas y simples, recorriendo paso a paso el camino de la generalización y desembocando finalmente en las grandes y maravillosas ideas que aparecen en esta teoría, las cuales resultan ser de capital importancia en estudios posteriores de matemáticas o de ciencia aplicada.

Éste es precisamente el espíritu que tiene la presente obra. El material aquí presentado tiene un carácter elemental (lo cual no significa que sea fácil) en cuanto que no pretende resaltar los aspectos más generales de las ideas expuestas. Sin embargo, se ha tratado siempre de explotar al máximo estas ideas con el fin de obtener en cada tema el mayor número de resultados que este tratamiento elemental permita.

ÁLGEBRA LINEAL

Atlacomulco 499-501, Fracc. Ind. San Andrés Atoto
53500 Naucalpan de Juárez, Edo. de México
Miembro de la Cámara Nacional de la Industria Editorial, Reg. Núm. 1890

ISBN 968-422-777-9

3456789012 ING-91 9087643215

Impreso en México Printed in Mexico

Esta obra se terminó de
imprimir en Septiembre de 1995 en
Programas Educativos, S.A. de C.V.
Calz. Chabacano Núm. 65-A
Col. Asturias
Delegación Cuauhtémoc
06850 México, D.F.

Se tiraron 500 ejemplares

Las matemáticas constituyen una actividad regida por las mismas reglas impuestas a las sinfonías de Beethoven, las pinturas de Da Vinci y las poesías de Homero. Así como las escalas, las leyes de la perspectiva y la métrica parecen carecer de vitalidad y ardor, podrá parecer que las reglas formales de las matemáticas son tediosas. Sin embargo, las matemáticas alcanzan pináculos tan elevados como los logrados por la imaginación de sus más osados exploradores. Y aquí se encierra, quizá, la última paradoja de la ciencia, puesto que en su prosaico tráfago, tanto la lógica como las matemáticas dejan atrás, frecuentemente, a su avanzada y muestran que el mundo de la razón pura es más extraño aún que el mundo de la fantasía pura.

E. Kasner/J. Newman
(*Matemáticas e imaginación*)

El álgebra lineal es una base que soporta otros estudios importantes en matemáticas (los cuales a su vez resultan ser centrales para los ingenieros, los físicos, etc.). En muchos casos importantes, la manera de actuar en este soporte es proporcionando un lenguaje adecuado y preciso para estudiar algunas otras ideas matemáticas. Un ejemplo de este hecho se encuentra en el estudio del cálculo de funciones del tipo $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, con n y m números naturales arbitrarios, en donde para establecer de una manera precisa el concepto de derivada de una función, es necesario hacerlo con el lenguaje del álgebra lineal.

De esta manera, cada libro sobre el tema presenta una manera distinta de construir la base-soporte del álgebra lineal. La elaboración que presenta este libro pretende ser simple pero sólida, de modo que, por una parte, el estudiante la pueda incorporar a sus conocimientos sin mucha dificultad y, por otra, pueda servirle adecuadamente como apoyo para estudios posteriores.

El lector podrá ver distribuidos a lo largo de la obra una buena cantidad de apéndices, tanto de secciones particulares de algún capítulo, como de capítulos mismos. Algunos de estos apéndices tienen la finalidad de exponer temas que, aun dentro del carácter elemental de toda la obra, son menos elementales que los temas que conforman propiamente el cuerpo principal del libro, pero que sin embargo, dependiendo de las intenciones de cada curso particular, podrían ser incluidos en un primer o en un segundo curso sobre álgebra lineal (tal es el caso de los apéndices I y II del capítulo 4, por ejemplo). Algunos otros apéndices no contienen material nuevo sobre álgebra lineal. En ellos, la intención es usar algunas ideas previamente estudiadas para aplicarlas a la solución de varios problemas "prácticos" (éste es el caso del apéndice de la sección 4 del capítulo 5, por ejemplo).

De cualquier modo, el conjunto de apéndices de este libro constituye un material opcional para un primer curso estándar y elemental sobre esta materia.

La secuencia que siguen los temas desarrollados a lo largo de los capítulos es tal que cada capítulo requiere, en general, del conocimiento del material discutido en los capítulos que le anteceden. En otras palabras, el libro presenta una estructura de exposición piramidal, teniendo la base en el capítulo 1 y en la cúspide el capítulo 7.

Se incluye un capítulo de conceptos generales (véase la introducción) que presentan análisis bastante concisos sobre tres temas importantes de cultura general en matemáticas: conjuntos (que se usa constantemente en todo el libro), relaciones de equivalencia (que se usa en los capítulos 3, 4 y 6) y funciones (uno de los conceptos más importantes de la matemática, que se usa, bajo diversas formas, en todos los capítulos).

En el capítulo 1 se estudian los sistemas de ecuaciones lineales y las matrices. En las secciones 1 y 2 se hace un estudio bastante elemental sobre sistemas de ecuaciones lineales. El interés principal es establecer una herramienta con la que se trabajará en el resto del libro. Es por eso que el énfasis en el tratamiento de este tema es en este momento de carácter operativo: se trata de ver *cómo* resolver sistemas de ecuaciones lineales. Al intentar contestar esta pregunta se producirán las primeras cuestiones teóricas interesantes alrededor de este tema, el cual es retomado en la sección 7 del capítulo 4, en donde, ya con un lenguaje propio del álgebra lineal (transformaciones lineales entre espacios vectoriales) se obtienen algunos otros resultados teóricos "no elementales" sobre sistemas de ecuaciones lineales y matrices. Las dos secciones restantes de este capítulo se dedican a estudiar el importante concepto de matriz, así como a establecer algunas de las propiedades más importantes que tienen estos "entes matemáticos". (En realidad, este concepto es introducido desde la sección 2 (subsección 2.1) en un intento de simplificación en la notación para el proceso de solución de sistemas de ecuaciones lineales.)

El capítulo 2 contiene un acercamiento elemental a la teoría de determinantes. Al igual que en el capítulo 1, no se pretende hacer un estudio exhaustivo de esta teoría, pues, por una parte, éste quedaría fuera de los alcances y objetivos de la obra, y, por otra parte, el interés en este capítulo sigue siendo desarrollar más herramientas que nos permitan navegar, con el menor número de dificultades técnicas, por los capítulos subsecuentes.

Podríamos decir que el capítulo 3 es el "primer" capítulo del libro, en el sentido de que en él se estudia la primera de las abstracciones fundamentales en el álgebra lineal: el concepto de espacio vectorial. A partir de este capítulo, y en lo que resta del libro, se puede ver el uso que se hace de las ideas y métodos desarrollados en los dos primeros capítulos con el fin de lograr una mejor comprensión de los puntos que involucra el estudio de esas maravillosas y potentes ideas que aparecen en el álgebra lineal. El sabor de la abstracción matemática, que apenas se esbozó en los capítulos 1 y 2, comienza a hacerse más palpable a partir de este capítulo. La manera como se pretende lograr que el estudiante entienda estas nuevas abstracciones matemáticas es mostrando una gran cantidad de ejemplos concretos (algunos de ellos numéricos) que ilustren cada uno de los conceptos introducidos. Es por eso que después de presentar una definición o de obtener un resultado, aparece, casi por regla general, la frase "por ejemplo . . ."

El capítulo 4 trata sobre las transformaciones lineales. La sospecha de que las transformaciones lineales y las matrices sean parientes muy cercanos trata de sembrarse desde la primera sección, en donde se da la definición y los primeros ejemplos de transformaciones lineales. Esta sospecha se convierte en una realidad en la sección 3. El lector interesado en ir más al fondo de esta relación entre transformaciones lineales y matrices puede estudiar el apéndice de la sección 5, en donde se muestra que el álgebra de las transformaciones lineales entre espacios vectoriales de dimensión finita es isomorfa al álgebra de matrices. En la sección 6 se retoma el análisis que sobre isomorfismos entre espacios vectoriales se hizo en la sección 6 (subsección 6.2) del capítulo 3. Esta sección contiene un apéndice que muestra cómo pueden ser usados los métodos del álgebra lineal para hallar antiderivadas de algunas funciones. Este "método" de integración no tiene valor alguno desde el punto de vista práctico, pero sí lo tiene desde el punto de vista de la belleza de las ideas involucradas. Se presentan, finalmente, dos apéndices en el capítulo 4. En el primero de ellos se estudian los funcionales lineales: se presenta el concepto de espacio dual a un espacio vectorial, el concepto de espacio bidual y su isomorfismo natural con el espacio del que procede, el concepto de transpuesta de una transformación lineal y, por último, se da una caracterización de los hiperplanos como núcleos (trasladados) de funcionales lineales. El apéndice II introduce el importante concepto de espacio vectorial cociente. Al lector con

intereses dentro de las matemáticas le es ampliamente recomendado el estudio de este apéndice, pues la construcción de estructuras cocientes es una de las herramientas más usadas en muchos otros estudios de matemáticas.

En el capítulo 5 se estudian de nuevo los espacios vectoriales, pero ahora con un nuevo ingrediente que les da una gran riqueza estructural: el producto interno. Este ingrediente permite, en muchas ocasiones, tener imágenes geométricas de los resultados obtenidos. La sección 4, que trata sobre bases ortonormales, contiene un apéndice sobre el método de los mínimos cuadrados, el cual es muy usado dentro de ciertos estudios de ingeniería y de ciencia aplicada en general.

El capítulo 6 es, sin duda, uno de los capítulos más importantes del libro. En él se estudia lo relacionado a los valores y vectores propios de transformaciones lineales y matrices. En este capítulo se da respuesta a una pregunta que fue planteada en la sección 4 del capítulo 4: dada una matriz cuadrada, ¿es ésta semejante a una matriz diagonal? La respuesta, como se ve en este capítulo, es que desgraciadamente no cualquier matriz es semejante a una matriz diagonal; esto es, no toda matriz es diagonalizable. Esta propiedad de diagonalización de una matriz es estudiada en la sección 3 usando el polinomio mínimo de la matriz (teorema 3.6). En esta sección también se establece el célebre teorema de Hamilton-Cayley. Se incluye un apéndice sobre teoría de gráficas. Ésta es una parte de la matemática contemporánea de gran riqueza y utilidad práctica. En este apéndice solamente se muestran algunos aspectos de la teoría que tienen conexión con el álgebra lineal.

En el capítulo 7 se estudian las formas bilineales y cuadráticas. En realidad, estos temas pertenecen ya a los terrenos del álgebra multilineal, pero, debido a las muy importantes aplicaciones que tienen las formas cuadráticas (como por ejemplo las presentadas en las secciones 3 y 4 de este capítulo) es que muchas veces se incluye su estudio en un curso introductorio de álgebra lineal.

El libro contiene más de 900 problemas propuestos para que el estudiante resuelva. En el apéndice se dan las respuestas, y en algunos casos, las soluciones explícitas de la mayoría de estos problemas. El papel que desempeña la resolución de éstos en la comprensión de la teoría discutida en el texto es, como en la mayoría de los libros de matemáticas, fundamental. El quedarse contemplando las definiciones y los teoremas que aparecen en el libro, seguramente no conducirá al estudiante a la asimilación de la teoría. Muchos de estos problemas son numéricos, de carácter operativo. A este respecto, es necesaria una aclaración: el álgebra lineal no es solamente un cúmulo de técnicas de cómo resolver sistemas de ecuaciones, cómo hallar bases de espacios vectoriales, cómo determinar los valores y vectores propios de una matriz, etc. Los problemas numéricos tienen muchas veces la intención de reforzar este aspecto operativo (que ciertamente ayuda a la comprensión de la materia). Sin embargo, un curso de álgebra lineal debe ir más lejos del carácter operativo, tratando de instalar en el estudiante una serie de conocimientos que, por una parte, le den *formación* dentro del tipo de razonamiento analítico propio de la matemática, y, por otra, le sirvan en futuros estudios de esta u otra parte del conocimiento científico. Es por eso que en algunos otros de los problemas propuestos se pide "demuestre que . . .". En ellos, se invita al estudiante a echar mano de la teoría previamente estudiada para establecer la validez del resultado indicado en el problema.

Se han clasificado los problemas propuestos en dos clases muy generales: los que no requieren más que de la aplicación directa de la teoría discutida en el texto (previamente asimilada, claro está) y los que presentan cierto grado de dificultad adicional, ya sea porque la habilidad operativa que demanda para su solución es alta, o porque quizás requiere de un "chispazo adicional" a la teoría discutida en el texto para demostrar un cierto resultado. Estos últimos han sido marcados con un pequeño círculo y, a su vez, se han subdividido en cuatro subclases; el número dentro del círculo indica el "grado de dificultad".

Por último, deseo hacer patente mi agradecimiento a las autoridades de la Universidad Panamericana que me brindaron su apoyo para la realización de esta obra. Agradezco también al Dr. Francisco González Acuña el haber usado como texto la primera versión de este libro en un curso que impartió en 1990; fruto de ello fueron una serie de valiosos comentarios que influyeron en la estructura de la versión definitiva de la obra que ahora se presenta. Agradezco también a la Srita. Irma Cuéllar Carmona quien realizó el trabajo mecanográfico original, a Alejandro López López y Ricardo Téllez Reyes, quienes trabajaron conmigo durante todo un semestre en la titánica labor de revisión de los originales y, muy especialmente, al matemático Alfonso Leal Guajardo, quien me ayudó a detectar algunos de los errores "más escondidos" que se encontraban todavía en la primera versión del libro, así como a presentar las soluciones de muchos de los problemas propuestos.

México, D.F., enero de 1991

Claudio Pita Ruiz
Escuela de Ingeniería
Universidad Panamericana

Contenido

INTRODUCCIÓN

Conceptos preliminares

La finalidad de este capítulo es proporcionar el material matemático básico con el que se trabajará en este libro. Los temas que se discutirán aquí son:

1) conjuntos
2) relaciones de equivalencia
3) funciones

El primero de estos temas es usado frecuentemente durante el desarrollo del libro. Los dos restantes se presentan pues, cuando se llegue a hacer uso de ellos, será necesario tener bien claros algunos detalles de los mismos.

No se pretende hacer una discusión exhaustiva del material que estos temas cubren. La discusión que se hará será más bien concisa, guiándose sobre todo por aquellos aspectos que llegarán a ser usados en algún momento en este libro. La demostración de algunos de los teoremas aquí presentados serán desarrollados como ejercicio por el lector.

A. CONJUNTOS

Como es usual en libros que no tratan la teoría de conjuntos desde un punto de vista avanzado, la idea de "conjunto" en este libro será aquella que intuitivamente da esa palabra, o bien, lo que intuitivamente se entiende por "familia" o "colección" de ciertos objetos. A estos objetos se le llamará *elementos* del conjunto. Es usual designar con letras mayúsculas a los conjuntos y con minúsculas a sus elementos.

La teoría de conjuntos es una de las partes "delicadas" de las matemáticas. Se advierte que este pequeño resumen de la teoría que aquí se presenta no es para especialistas. El lector interesado en ver con profundidad esta teoría, puede consultar —como primer paso— el libro de P. R. Halmos, *Naive Set Theory,* Van Nostrand, Princeton, N. J. (1960).

Cuando es posible exhibir de alguna manera los elementos de un conjunto, dígase $x_1, x_2, \ldots$, se escribe

$$\{x_1, x_2, \ldots\}$$

para denotar al conjunto cuyos elementos son $x_1, x_2, \ldots$ Sin embargo, muchas veces no es posible mostrar explícitamente los elementos de un conjunto. En tal caso se recurre a exhibir alguna propiedad compartida por todos los elementos de él. Dígase que P es tal propiedad. Entonces se escribe

$$\{x \mid P\}$$

para denotar al conjunto cuyos elementos son las x que satisfacen la propiedad P. Se dice que un conjunto es finito (infinito) si el número de sus elementos es finito (infinito, respectivamente).

Algunos conjuntos importantes que se usarán en este libro (todos ellos infinitos)

$\mathbf{N} = \{1, 2, 3, \ldots\}$ el conjunto de números naturales
$\mathbf{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ el conjunto de números enteros
$\mathbf{R}$ = el conjunto de números reales
$\mathbf{C} = \{a + ib \mid a, b \in R\}$ el conjunto de números complejos

Si x es un elemento del conjunto A se escribe

$$x \in A$$

En caso contrario se escribe

$$x \notin A$$

Por ejemplo, $3 \in \mathbf{N}$ y $\sqrt{-1} \notin \mathbf{R}$.

A un conjunto que carece de elementos se le llamará *conjunto vacío* y se denotará por $\emptyset$.

Se dice que los conjuntos A y B son *iguales*, lo cual se representa como $A = B$, si ellos constan exactamente de los mismos elementos. En caso contrario se dice que A y B son diferentes y se escribe $A \neq B$. Por ejemplo si

$$\begin{aligned} A &= \{2, 4, 6, \ldots\} \\ B &= \{x \mid x = 2k, k \in \mathbf{N}\} \\ C &= \{x \mid x = 2k - 1, k \in \mathbf{N}\} \end{aligned} \qquad (1)$$

se tiene que $A = B \neq C$.

Se dice que el conjunto A es un *subconjunto* del conjunto B, o bien que el conjunto B *contiene* al conjunto A, lo cual se escribe como $A \subseteq B$, si todo elemento del conjunto A es también elemento del conjunto B. Es decir $A \subseteq B$ si

$$x \in A \Rightarrow x \in B$$

Por ejemplo, todo conjunto es subconjunto de sí mismo, y si A es un conjunto cualquiera, el conjunto vacío es un subconjunto de A.

La relación de igualdad entre conjuntos puede ser establecida en términos de contención entre los conjuntos: se dice que el conjunto A es igual al conjunto B si

$$A \subseteq B \quad \text{y} \quad B \subseteq A$$

Probar este par de contenciones es el procedimiento normal que se sigue cuando se quiere demostrar que los conjuntos A y B son iguales.

Cuando un conjunto A es subconjunto del conjunto B, pero $A \neq B$, se dice que A es un *subconjunto propio* de B, o bien que el conjunto B *contiene propiamente* al conjunto A. Para enfatizar este hecho, en ocasiones se escribe $A \subsetneq B$.

Dados los conjuntos A y B, se define el conjunto *unión de A y B*, denotado por $A \cup B$, como aquel cuyos elementos son los elementos que pertenecen a A o los que pertenecen a B (o a ambos). Es decir

$$A \cup B = \{x \mid x \in A \quad \text{o} \quad x \in B\}$$

Por ejemplo, si A y C son los conjuntos definidos en (1), se tiene que

$$\mathbf{N} = A \cup C$$

Las principales propiedades de la unión de conjuntos de resumen en el siguiente teorema:

TEOREMA 1.1

Si A, B y C son conjuntos cualesquiera, entonces

1) $A \cup B = B \cup A$
2) $A \cup (B \cup C) = (A \cup B) \cup C$
3) $A \cup A = A$
4) $A \cup \emptyset = A$
5) $A \subseteq B \Rightarrow A \cup B = B$

Dados los conjuntos A y B, defina el conjunto *intersección* de A y B, denotado por $A \cap B$, como aquel cuyos elementos pertenecen a A y a B simultáneamente. Es decir

$$A \cap B = \{x \mid x \in A \quad \text{y} \quad x \in B\}$$

Si $A \cap B = \emptyset$ (esto es, el conjunto A no tiene elementos en común con el conjunto B), diga que los conjuntos A y B son *disjuntos* (o ajenos). Por ejemplo, los conjuntos A y C de (1) son disjuntos.

Las principales propiedades de la intersección de conjuntos se resumen a continuación.

TEOREMA 1.2

Si A, B y C son conjuntos cualesquiera, se tiene

1) $A \cap B = B \cap A$
2) $A \cap (B \cap C) = (A \cap B) \cap C$
3) $A \cap A = A$
4) $A \cap \emptyset = \emptyset$
5) $A \subseteq B \Rightarrow A \cap B = A$

Dados los conjuntos A y B, defina el conjunto *diferencia* entre A y B, denotado por $A \setminus B$,* como el conjunto cuyos elementos son los elementos del conjunto A que no pertenecen al conjunto B. Es decir

$$A \setminus B = \{x \mid x \in A \quad \text{y} \quad x \notin B\}$$

Por ejemplo si A es el conjunto definido en (1) y B es el conjunto

$$B = \{x \mid x = 4k, \ k \in \mathbb{N}\}$$

se tiene:

$$A \setminus B = \{2, 6, 10, \ldots\}$$

TEOREMA 1.3

> Si A y B son conjuntos cualesquiera, se tiene
>
> 1) $A \setminus \emptyset = A$
> 2) $A \cap B = \emptyset \Rightarrow A \setminus B = A$
> 3) $A \subseteq B \Rightarrow A \setminus B = \emptyset$

Si A y B son dos conjuntos cualesquiera, defina el *producto cartesiano* de A y B, el cual se denota como $A \times B$, como el conjunto de todas las parejas ordenadas (a, b) cuyo primer elemento pertenece a A y el segundo a B. Es decir

$$A \times B = \{(a, b \mid a \in A, b \in B\}$$

Por ejemplo si $A = \{1, 2, 3\}$ y $B = \{1, 2\}$ se tiene

$$A \times B = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 1), (3, 2)\}$$

Más generalmente, con los n conjuntos $A_1, A_2, \ldots, A_n$ puede formar su producto cartesiano $A_1 \times A_2 \times \ldots \times A_n$, el cual es (por definición) el conjunto de n-adas ordenadas $(a_1, a_2, \ldots a_n)$ en donde $a_i \in A_i$, $i = 1, 2, \ldots, n$. En el caso en el que $A_i = A$, $i = 1, 2, \ldots, n$, se escribe

$$\underbrace{A \times A \times \ldots \times A}_{n \text{ veces}} = A^n$$

En este libro se estudiará (en el capítulo 3) el conjunto R^n el cual es entonces

$$R^n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in R, \ i = 1, 2, \ldots, n\}$$

B. RELACIONES DE EQUIVALENCIA

Una *relación binaria* en un conjunto X es un símbolo matemático o una afirmación verbal, que se denota por R, tal que para cada par de elementos x, y en X, la aserción xRy tiene completo sentido y puede asignársele sin ambigüedad el valor de verdad o falsedad; es decir, dados $x, y \in X$, la aserción xRy es verdadera o falsa. Cuando

*Nota: La diferencia de A y B se denota a veces por $A - B$, o bien por $A \sim B$.

xRy es verdadera, se dice que el elemento x está relacionado por medio de la relación binaria R con el elemento y, y se escribe simplemente xRy. Caso contrario, es decir cuando xRy es falsa, se dice que x no está relacionado con y, y se escribe $x\not{R}y$. Por ejemplo, si X es el conjunto de los números enteros Z, la relación $<$ ("menor que") es una relación binaria en X. Así se tiene que $-1 < 2$ y $-1 \not< -3$.

En esta sección es de especial interés estudiar un tipo concreto de relaciones binarias llamadas relaciones de equivalencia.

Una *relación de equivalencia* en un conjunto X es una relación binaria R en X que cumple las siguientes propiedades:

1. **reflexividad:** xRx. $\forall \ x \in X$.
2. **simetría:** $xRy \Rightarrow yRx$, $\forall \ x, y \in X$
3. **transitividad:** xRy, $yRz \Rightarrow xRz$, $\forall \ x, y, z \in X$

Se usará frecuentemente el símbolo $\sim$ para denotar a la relación R.

Sea $\sim$ una relación de equivalencia en un conjunto X. Para cada elemento $x \in X$ se define *la clase de equivalencia de* x denotada por $[x]$,* como el conjunto de todos los elementos de X que están relacionados con x por medio de la relación $\sim$. O sea

$$[x] = \{y \in X \mid y \sim x\}$$

Si X es un conjunto (no vacío) cualquiera, el ejemplo más simple de relación de equivalencia se encuentra si se define la relación $\sim$ en X como

$$x \sim y \Leftrightarrow x = y$$

O sea que $\sim$ es la relación de igualdad entre los elementos de X. Es claro que esta relación es **reflexiva, simétrica** y **transitiva**. La clase de equivalencia de x consta de un solo elemento pues

$$[x] = \{y \in X \mid y = x\} = \{x\}$$

La relación $<$ en el conjunto de números enteros no es una relación de equivalencia, pues no es **reflexiva** ($x \not< x$) ni **simétrica** (si $x < y$ no es cierto que $y < x$). Esta relación binaria cumple sólo la propiedad de **transitividad**. Del mismo modo, la relación $\leq$ en el conjunto de números enteros tampoco es una relación de equivalencia, pues no es **simétrica** (si $x \leq y$, no es cierto en general que $y \leq x$, a menos que $x = y$). Obsérvese sin embargo que esta relación sí es reflexiva, a diferencia de la primera. También es **transitiva**.

Sea X el conjunto de mexicanos. En X se establece la relación binaria $\sim$ entre sus elementos como

$$x \sim y \Leftrightarrow x \text{ y } y \text{ nacieron en el mismo estado}$$

*Esta notación para una clase de equivalencia no es universal. De hecho, en el libro se llegarán a usar notaciones distintas a ésta.

Se verifica fácilmente que ésta es una relación de equivalencia en X. La clase de equivalencia de un elemento $x \in X$, está constituida por todos aquellos mexicanos que nacieron en el mismo estado que x. Obsérvese que si x y y nacieron en el mismo estado, entonces $[x] = [y]$, mientras que si x y y no nacieron en el mismo estado, se tiene $[x] \cap [y] = \emptyset$. Este es un hecho general que se probará más adelante.

Véase un último ejemplo. Sea X el conjunto de los números enteros **Z** y n un número natural. En X se define la relación binaria ~ entre sus elementos como

$$x \sim y \Leftrightarrow \exists\, k \in \mathbf{Z} \text{ tal que } x - y = kn$$

Es decir, los enteros x, y están relacionados por medio de ~, si la diferencia $x - y$ es un múltiplo de n. Verifíquese que ésta es una relación de equivalencia en **Z**.

Para cualquier entero x se tiene que $x \sim x$, pues $x - x = 0 = 0 \cdot n$. Esto muestra la propiedad reflexiva de la relación ~. Supóngase ahora que $x \sim y$. Existe entonces $k \in \mathbf{Z}$ tal que $x - y = kn$. Pero en tal caso, $y - x = (-k)n$, $-k \in \mathbf{Z}$, es decir que $y \sim x$, lo que muestra la propiedad simétrica de la relación ~. Finalmente, supóngase que $x \sim y$ y $y \sim z$. Existen enteros k_1 y k_2 tales que $x - y = k_1 n$, $y - z = k_2 n$. Entonces $x - z = x - y + y - z = k_1 n + k_2 n = (k_1 + k_2)n$, $k_1 + k_2 \in \mathbf{Z}$, es decir que $x \sim z$, lo que muestra que la relación ~ posee la propiedad transitiva. Se concluye pues que ~ es una relación de equivalencia.

La clase de equivalencia de un elemento $x \in \mathbf{Z}$ es entonces

$$[x] = \{y \in \mathbf{Z} \mid y \sim x\} = \{y \in \mathbf{Z} \mid y - x = kn,\ k \in \mathbf{Z}\} = \{y \in \mathbf{Z} \mid y = x + kn,\ k \in \mathbf{Z}\}$$

Por ejemplo, si $n = 4$ se tiene

$$[0] = \{y \in \mathbf{Z}\ \ y = 4k,\ k \in \mathbf{Z}\} = \{0, \pm 4, \pm 8, \ldots\}$$
$$[1] = \{y \in \mathbf{Z}\ \ y = 1 + 4k,\ k \in \mathbf{Z}\} = \{\ldots, -11, -7, -3, 1, 5, 9, \ldots\}$$
$$[2] = \{y \in \mathbf{Z}\ \ y = 2 + 4k,\ k \in \mathbf{Z}\} = \{\pm 2, \pm 6, \pm 10, \ldots\}$$
$$[3] = \{y \in \mathbf{Z}\ \ y = 3 + 4k,\ k \in \mathbf{Z}\} = \{\ldots, -9, -5, -1, 3, 7, 11, \ldots\}$$
$$[4] = \{y \in \mathbf{Z}\ \ y = 4 + 4k,\ k \in \mathbf{Z}\} = \{y \in \mathbf{Z}\ \ y = 4k,\ k \in \mathbf{Z}\} = [0]$$

En este caso existen solamente 4 clases de equivalencia distintas, a saber [0], [1], [2], [3].

El resultado principal de esta sección se encuentra en el siguiente teorema:

TEOREMA 2.1

> Sea X un conjunto y sea ~ una relación de equivalencia en X. Las distintas clases de equivalencia obtenidas en esta relación proporcionan una descomposición de X como una unión de subconjuntos ajenos. Recíprocamente, dada una descomposición de X como unión de subconjuntos ajenos no vacíos, es posible definir una relación de equivalencia ~ en X tal que estos subconjuntos sean precisamente las clases de equivalencia obtenidas de la relación ~.

DEMOSTRACIÓN Obsérvese primeramente que

$$X = \bigsqcup_{x \in X} [x]$$

Se tiene que probar entonces que dadas dos clases de equivalencia $[x]$ y $[y]$ se tiene una de las siguientes posibilidades:

1. $[x] = [y]$
2. $[x] \cap [y] = \emptyset$

Supóngase que $[x] \cap [y] \neq \emptyset$. Se probará entonces que en este caso las clases $[x]$ y $[y]$ coinciden.

Sea $z \in [x] \cap [y]$. Entonces

$$z \sim x \tag{1}$$

y

$$z \sim y \tag{2}$$

Tómese un elemento a en $[x]$. Entonces $a \sim x$; pero por (1), $x \sim z$ (pues ~ es simétrica) y por tanto $a \sim z$ (pues ~ es transitiva). Por (2), $a \sim y$ (pues ~ es transitiva) lo que quiere decir entonces que $a \in [y]$. Entonces $[x] \subseteq [y]$.

Un argumento similar nos conduce a la otra contención $[y] \subseteq [x]$. Se concluye pues que $[x] = [y]$.

Esto muestra la primera parte del teorema.

Supóngase ahora que X es una unión de subconjuntos ajenos no vacíos. Defínase en X la relación ~ como

$$x \sim y \Leftrightarrow x \text{ y } y \text{ pertenecen al mismo subconjunto}$$

Se verifica de inmediato que ésta es una relación de equivalencia y que las clases de equivalencia son precisamente los subconjuntos en que estaba descompuesto inicialmente X.

Q.E.D.

Cuando en un conjunto X se establece una relación de equivalencia entre sus elementos, se dice que "X queda partido —o dividido— en clases de equivalencia", indicando con esto el hecho de que X puede ser contemplado como la unión disjunta de las clases de equivalencia de sus elementos. Este hecho resulta ser de capital importancia dentro de algunas situaciones concretas en matemáticas. Por ejemplo, esto permite hacer una *identificación* de todos los elementos de X que pertenezcan a una misma clase de equivalencia: identificándose cada elemento $x \in X$ con la clase de equivalencia a la que pertenece, y de esta manera dos elementos que pertenezcan a la misma clase serán, bajo esta perspectiva, indistinguibles (esta situación se presenta en la discusión sobre isomorfismos de espacios vectoriales; subsección 6.2 del capítulo 3).

Otra situación que se encontrará también en este libro, relacionada con la partición de un conjunto X en clases de equivalencia, será la de procurar "representantes adecuados" en una cierta clase de equivalencia: supóngase que a cada elemento de un conjunto X se le puede asociar una y solamente una clase de equivalencia de otro conjunto Y. Habrá que preguntarse entonces si dentro de la clase de equivalencia que le corresponde a un cierto elemento $x \in X$, existe un elemento (un representante de la clase) con ciertas características. Esta es de hecho la situación abstracta de la que trata el capítulo 6 de este libro. Por último, quizás la situación más importante relacionada con partición de un conjunto X en clases de equivalencia, desde el punto de vista del álgebra, es la siguiente: supóngase que el conjunto X posee una cierta estructura algebraica (grupo, anillo, espacio vectorial, módulo). En X se establece una relación de equivalencia entre sus elementos de modo que X queda dividido en clases de equivalencia. Denótese por $\tilde{X}$ al conjunto de todas las clases de equivalencia así obtenidas. Habrá que preguntarse si el conjunto $\tilde{X}$ posee también la misma estructura algebraica de X. En tal caso se dice que $\tilde{X}$ posee la *estructura cociente* y se escribe $\tilde{X} = X/\sim$, en donde ~ es la relación de equivalencia que se definió en X. En el apéndice II del capítulo 4, se presenta una discusión detallada de cómo es que se obtiene la estructura cociente en un espacio vectorial.

C. FUNCIONES

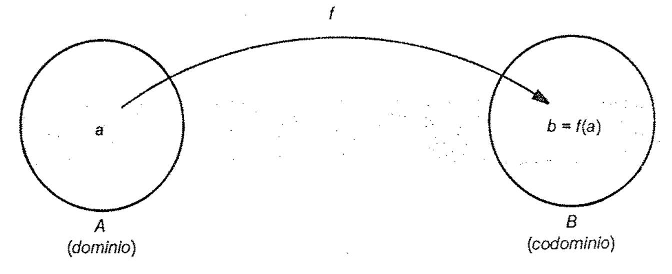
Sean A y B dos conjuntos no vacíos. Se dice que f es una *función* del conjunto A al conjunto B, lo cual se escribe como

$$f: A \rightarrow B$$

si a cada elemento $a \in A$ le corresponde un único elemento $b \in B$, el cual se denota por $b = f(a)$. Se dice en tal caso que b es la imagen de a bajo f, y que el elemento $a \in A$ es el *argumento* de la función f.

Al conjunto A se le llama *dominio* de la función f, y al conjunto B se le llama *codominio* de la función f.

Esquemáticamente tenemos



Establecer entonces una función f consiste en dar

- su dominio (el conjunto A)
- su codominio (el conjunto B)
- la regla de correspondencia por medio de la cual a cada elemento a del conjunto A se le asocia el elemento $b = f(a)$ del conjunto B.

En los primeros cursos de cálculo se estudian funciones para las cuales el conjunto B es el conjunto de números reales $\boldsymbol{R}$ y el conjunto A es un subconjunto de $\boldsymbol{R}$; un intervalo. Es decir, funciones del tipo $f: A \subseteq \boldsymbol{R} \rightarrow \boldsymbol{R}$, llamadas "funciones reales de variable real". En esta sección no se insistirá demasiado sobre este tipo de funciones pues aquellas funciones que aparecen en el estudio del álgebra lineal son de un tipo más general.

Véanse ahora algunos ejemplos.

Sea A un conjunto no vacío cualquiera. Sea $f: A \rightarrow A$ dada por $f(a) = a$, $\forall\, a \in A$. Es claro que f es una función del conjunto A en sí mismo. Esta función es llamada *función identidad* (en A) y se denota por Id_A.

Sean ahora A y B dos conjuntos no vacíos cualesquiera. Tómese un elemento fijo $b_o \in B$. Si se considera $f: A \rightarrow B$ como $f(a) = b_o$, $\forall\, a \in A$, es claro que se tendrá también una función del conjunto A al conjunto B. Esta función es llamada *función constante*.

Algunos otros ejemplos de funciones son:

$f: N \rightarrow N$, $f(x) = 2x$
$f: \boldsymbol{R} \rightarrow \boldsymbol{R}$, $f(x) = x^2 + 5x + \pi$
$f: \boldsymbol{Z} \rightarrow N$, $f(x) = x^2$
$f: \boldsymbol{R}^2 \rightarrow \boldsymbol{R}$, $f(x, y) = x + y$
$f: N \rightarrow N^2$, $f(x) = (x, 2x)$
$f: \boldsymbol{Z}^2 \rightarrow \boldsymbol{Z}^3$, $f(x, y) = (x + y, x - y, 2x + 2y)$, etcétera.

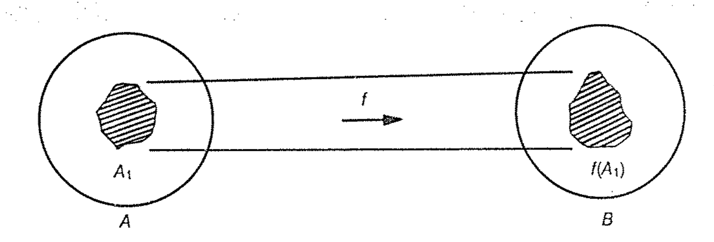
Se dice que las funciones $f_1: A_1 \rightarrow B_1$ y $f_2: A_2 \rightarrow B_2$ son *iguales*, lo cual se escribe como $f_1 = f_2$, si $A_1 = A_2$, $B_1 = B_2$ y $f_1(a) = f_2(a)$, $\forall\, a \in A_1(=A_2)$.

Sea ahora $f: A \rightarrow B$ una función del conjunto A al conjunto B. Si A_1 es un subconjunto de A, se dice que el conjunto $f(A_1)$, subconjunto de B, dado por

$$f(A_1) = \{f(a) \mid a \in A_1\} \subseteq B$$

es la *imagen* (directa) *de* A_1 *bajo* f.

Esquemáticamente



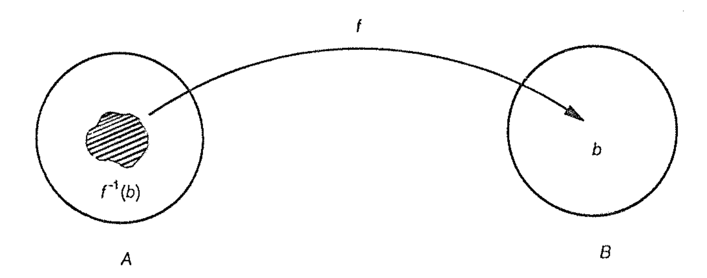
Si B_1 es un subconjunto de B, al conjunto $f^{-1}(B_1)$, subconjunto de A, dado por

$$f^{-1}(B_1) = \{a \in A \mid f(a) \in B_1\} \subseteq A$$

se le llama *imagen inversa de* B_1. Si B_1 está constituido por un solo elemento, dígase b, entonces se escribe

$$f^{-1}(b) = \{a \in A \mid f(a) = b\} \subseteq A$$

y se dice que este subconjunto de A es la imagen inversa de b.



Obsérvese que $f^{-1}(b)$ puede ser vacío.

Por ejemplo, si se considera la función f: $R \to R$, $f(x) = x^2$, se tiene

$$f^{-1}(-1) = \emptyset$$
$$f^{-1}(0) = \{0\}$$
$$f^{-1}(1) = \{1, -1\}$$

Más generalmente:

$$f^{-1}(y) = \emptyset \text{ si } y < 0$$
$$f^{-1}(0) = 0$$
$$f^{-1}(y) = \{\sqrt{y}, -\sqrt{y}\} \text{ si } y > 0$$

Al conjunto de elementos $b \in B$ para los cuales $f^{-1}(b)$ no es vacío, se le llama *rango* de la función f.* Por ejemplo, para la misma función f: $R \to R$, $f(x) = x^2$, se tiene que

$$\text{rango de } f = \{y \in R \mid y \geq 0\}$$

Si para la función f: $A \to B$ se cumple que

$$\text{rango de } f = B$$

dígase que f es una función *sobreyectiva* (o *suprayectiva*). Si ocurre que para cualesquiera $a_1, a_2 \in A$, tales que $a_1 \neq a_2$, se tiene $f(a_1) \neq f(a_2)$, se dice que f es una *función inyectiva*. Equivalentemente, la función f: $A \to B$ es *inyectiva* si

*En el álgebra lineal la palabra *rango* aparece designando un concepto distinto al aquí presentado. En su momento (sección 2, capítulo 5) se hará de nuevo la aclaración correspondiente.

$$f(a_1) = f(a_2) \Rightarrow a_1 = a_2$$

Si una función es a la vez *inyectiva* y *sobreyectiva*, se dice que es una *función biyectiva*.

Por ejemplo, la función f: $N \to N$, $f(x) = 2x$ es *inyectiva* pero no *sobreyectiva*. La función f: $Z \setminus \{0\} \to N$, $f(x) = |x|$, es *sobreyectiva* pero no *inyectiva*. La función f: $R \to R$, $f(x) = x^4$ no es ni *inyectiva* ni *sobreyectiva*. Por último, la función f: $R \to R$, $f(x) = x^3$ es *inyectiva* y *sobreyectiva*; es por tanto una función *biyectiva*.

Sea ahora f: $A \to B$ una función del conjunto A al conjunto B y g: $B \to C$ una función del conjunto B al conjunto C. Defínase la *función composición de g con f*, denotada por $g \circ f$, como la función

$$g \circ f: A \to C$$
$$(g \circ f)(a) = g(f(a))$$

Esquemáticamente

$c = (g \circ f)(a) = g(f(a)) = g(b)$

Obsérvese que aunque se pueda establecer la composición de $g \circ f$, puede ocurrir que la composición $f \circ g$ no esté definida. Aún más, si ambas composiciones $g \circ f$ y $f \circ g$ estuvieran definidas, no es cierto en general que $g \circ f = f \circ g$. Debido a esto, se dice que la composición de funciones es una operación no conmutativa.

Por ejemplo, si se tienen las funciones

$$f: R^2 \to R^3,\ f(x, y) = (x + y, x - y, 1)$$
$$g: R^3 \to R,\ g(x, y, z) = x + y + z$$

se puede formar la composición $g \circ f$: $R^3 \to R$. La cual está dada por

$$(g \circ f)(x, y, z) = g(f(x, y, z)) = g(x + y, x - y, 1) = x + y + x - y + 1 = 2x + 1$$

En este caso *no* se puede formar la composición $f \circ g$ (¿por qué?).

Por otro lado, si se consideran las funciones

$$f: R \to R\ \ f(x) = x^2$$
$$g: R \to R\ \ g(x) = 2x$$

se tiene que

$$g{\circ}f\colon R \to R$$
$$(g{\circ}f)(x) = g(f(x)) = g(x^2) = 2x^2$$

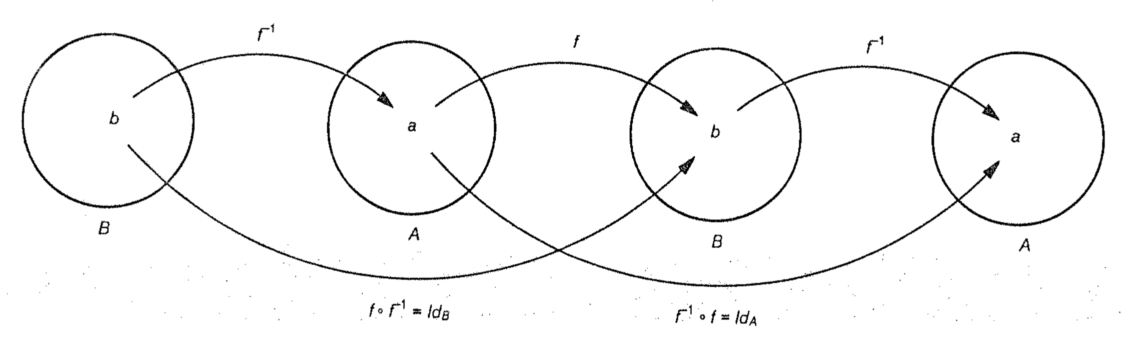
y

$$f{\circ}g\colon R \to R$$
$$(f{\circ}g)(x) = f(g(x)) = f(2x) = (2x)^2 = 4x^2$$

Se tiene pues que $g{\circ}f \neq f{\circ}g$.

Se dice que la función $f\colon A \to B$ *tiene inversa*, si existe una función $f^{-1}\colon B \to A$ (llamada *inversa de* f) tal que

$$f^{-1}{\circ} f = Id_A$$
$$f{\circ}f^{-1} = Id_B$$



Un resultado importante, que da condiciones equivalentes al hecho de que una función tenga inversa, está contenido en el siguiente teorema:

TEOREMA 3.1

> La función $f\colon A \to B$ tiene inversa si y solamente si es una función biyectiva.

DEMOSTRACIÓN Supóngase primeramente que la función $f\colon A \to B$ tiene inversa.

Véase que f es sobreyectiva: dado $b \in B$ se tiene

$$b = Id_B(b) = (f{\circ}f^{-1})(b) = f(f^{-1}(b))$$

Si $a = f^{-1}(b)$, hemos mostrado que $f(a) = b$. Entonces $b \in$ rango de f. Es decir, rango de $f = B$. Por tanto, f es *sobreyectiva*.

Véase ahora que f es *inyectiva*: sean $a_1, a_2 \in A$ tales que

$$f(a_1) = f(a_2)$$

Entonces

$$f^{-1}(f(a_1)) = f^{-1}(f(a_2))$$

Como $f^{-1}{\circ}f = Id_A$, se concluye que $a_1 = a_2$, lo que muestra que f es *inyectiva*.

Entonces f es *biyectiva*, como se quería.

Supóngase ahora que f es una función *biyectiva*. Se quiere ver que f tiene inversa. Para un elemento $b \in B$, existe un elemento $a \in A$ tal que $f(a) = b$ (pues f es *sobreyectiva*). Este elemento es único (pues f es inyectiva). Defina la función

$$f^{-1}\colon B \to A$$

como $f^{-1}(b) = a$. Es inmediato verificar que $f^{-1}{\circ}f = Id_A$ y $f{\circ}f^{-1} = Id_B$. Esto muestra que f tiene inversa.

Q.E.D.

Por ejemplo, si $f\colon R \to R$ es la función $f(x) = x^3$, se tiene que f tiene inversa, pues es una función *biyectiva*.

La función $f^{-1}\colon R \to R$, inversa de f es $f^{-1}(x) = \sqrt[3]{x}$.

EJERCICIOS (INTRODUCCIÓN)

1. Sean los conjuntos

$A = \{e,s,t,u,d,i,a,r\}$
$B = \{m,a,t,e,m,á,t,i,c,a,s\}$
$C = \{e,s\}$
$D = \{f,a,s,c,i,n,a,n,t,e\}$

a) Describa los conjuntos $A \cup B, A \cup D, A \cap B, A \cap D$.
b) Compruebe que C es subconjunto de A, B y D
c) Demuestre que $A \cap D \subset B \cap D$
d) Compruebe que $A \setminus D = A \setminus B$
e) Demuestre que $\{f,r,a,n,c,é,s\} \subset A \cup D$
f) Compruebe que $\{t,í,a\}$ es subconjunto de A, B y D

2. Sea A un conjunto no vacío y a un elemento de A. Sea $B = \{a\}$. Diga cuáles de las siguientes expresiones son correctas, cuáles no, o bien, sobre cuáles de ellas no se puede decidir (por falta de información). Justifique su respuesta.

a) $a \in B$	f) $A \subset B$	k) $A \cap B = a$
b) $a \in A$	g) $B \subset A$	l) $a \in A \cup B$
c) $B \in A$	h) $B \subset \{a\}$	m) $\{a\} \subset A \cap B$
d) $\{a\} \in B$	i) $\{a\} \in A$	n) $\{a\} \in A \cap B$
e) $\{a\} \subset B$	j) $\{a\} \subset A$	o) $A \cup B = \{a\}$

3. Sean los conjuntos $A = \{0,1\}$, $B = \{1,0\}$, $C = \{0, \{1\}\}$, $D = \{\{0\}, 1\}$, $E = \{\{0\}, \{0,1\}\}$, $F = \{\{0\}, \{1,0\}\}$. Manifieste cuáles de las siguientes expresiones son correctas y

cuáles no. Justifique su respuesta.

a) $A \subset B$	f) $E \subset F$	k) $E \cap F = \{1,0\}$
b) $A = B$	g) $F \subset E$	l) $E \cup F = \{0\} \cup A$
c) $A \subset C$	h) $A \in E$	m) $E \cap C = \{0\}$
d) $A \subset D$	i) $A \subset E$	n) $E \cap D = \{0\}$
e) $A = D$	j) $A \cap B = A$	o) $B \subset F$

4. Sea $A = \{\emptyset, \{\emptyset\}\}$ ($\emptyset$ es el conjunto vacío). Diga cuáles de las siguientes expresiones son correctas y cuáles no. Pruebe su respuesta.

a) A es el conjunto vacío	f) $A \cap \{\emptyset\} = \{\emptyset\}$
b) $\emptyset \subset A$	g) $A \cap \emptyset = \emptyset$
c) $\emptyset \in A$	h) $A \cup \{\emptyset\} = A$
d) $\{\emptyset\} \in A$	i) $A \cap \{\emptyset\} = \emptyset$
e) $\{\emptyset\} \subset A$	j) $A \cap \emptyset = \{\emptyset\}$

5. Sean a y b dos números cualesquiera y sea
$$A = \{a,b, \{a\}, \{b\}\}$$
¿Cierto o falso?: el conjunto A tiene sólo dos elementos distintos.
(*Sugerencia*: considere los dos casos $a = b$ y $a \neq b$.)
6. Con los números 0 y 1 construya un conjunto que tenga 5 elementos distintos.
7. Sea $A = \{xy \mid x, y \in N\}$. Demuestre que $A = N$.
8. Sea $B = \{x + y \mid x, y \in Z\}$. Compruebe que $B = Z$.
9. Sean A, B y C tres conjuntos tales que
$$A \subset B \subset C \subset A$$
Demuestre que $A = B = C$. Más generalmente, si $A_1, A_2, \ldots, A_n$ son n conjuntos tales que
$$A_1 \subset A_2 \subset \ldots \subset A_{n-1} \subset A_n \subset A_1$$
pruebe entonces que $A_1 = A_2 = \ldots = A_n$.
10. Demuestre el teorema 1.1.
11. Demuestre el teorema 1.2.
12. Demuestre el teorema 1.3.
13. Sean A, B y C tres conjuntos cualesquiera. Demuestre las leyes de De Morgan
    a) $C \setminus (A \cup B) = (C \setminus A) \cap (C \setminus B)$
    b) $C \setminus (A \cap B) = (C \setminus A) \cup (C \setminus B)$
14. Dado un conjunto A, se define su *conjunto potencia*, denotado por $p(A)$ como el conjunto formado por todos los subconjuntos de A. Es decir
$$p(A) = \{B | B \subset A\}$$
Por ejemplo, si $A = \{1,2\}$ entonces
$$p(A) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$$
    a) Si $A = \emptyset$, describa $p(A)$
    b) Si $A = \{1,2,3\}$, describa $p(A)$
    c) Si $A = \{1,2\}$, describa $p(p(A))$

① 15. Suponga que el conjunto A tiene n elementos. Demuestre que el conjunto $p(A)$ tiene
$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$
subconjuntos de A de k elementos cada uno. Concluya entonces que $p(A)$ tiene 2^n elementos.
① 16. Sea A un conjunto cualquiera y a un elemento de él. ¿Puede ocurrir que $a \in p(A)$? Explique.
17. Si los conjuntos A y B son iguales, claramente se tiene $p(A) = p(B)$. ¿Es verdadera la afirmación recíproca?
① 18. Se define la *diferencia simétrica* de los conjuntos A y B denotada por $A \Delta B$, como
$$A \Delta B = \{x \mid x \in A \text{ o } x \in B \text{ y } x \notin A \cap B\}$$
Demuestre que
$$A \Delta B = (A \setminus B) \cup (B \setminus A)$$
19. Sean A, B, C y D conjuntos cualesquiera. Compruebe que
    a) $A \Delta (B \Delta C) = (A \Delta B) \Delta C$
    b) $A \Delta A = \emptyset$
    c) $A \Delta \emptyset = A$
    d) $A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C)$
    e) $(A \cup B) \Delta (C \cup D) \subset (A \Delta C) \cup (B \Delta D)$
20. Dé una interpretación geométrica de los conjuntos $R \times R = R^2$ y $R \times R \times R = R^3$
21. En el plano xy describa geométricamente el conjunto $N \times N$. Haga lo mismo para el conjunto $\mathbf{Z} \times \mathbf{Z}$.
22. Sea A un conjunto cualquiera. ¿Qué es el conjunto $A \times \emptyset$?
23. Suponga que para dos conjuntos A y B se tiene $A \times B = \emptyset$. ¿Qué puede decir de estos conjuntos?
24. a) Suponga que $A \subset B$ y $C \subset D$. Demuestre entonces que $A \times C \subset B \times D$.
    b) Piense que para los conjuntos A, B, C y D se tiene $A \times C \subset B \times D$. ¿Es posible concluir que $A \subset B$ y $C \subset D$?

---

25. Sea X el conjunto de todos los seres humanos. En X defina la relación ~ entre sus elementos como
    a) $x \sim y$ si x es hermano de y
    b) $x \sim y$ si x es padre de y
    c) $x \sim y$ si x es amigo de y
    d) $x \sim y$ si x nació en la misma ciudad en que nació y
    e) $x \sim y$ si x vive en la misma casa que y
    f) $x \sim y$ si x habla la misma lengua que y
Diga en cada caso si ~ es una relación de equivalencia en X. En el caso de que lo sea, describa las clases de equivalencia correspondientes.
26. Considere la relación ~ en el conjunto $\mathbf{Z}$ definida como
$$x \sim y \Leftrightarrow \exists\, k \in Z \text{ tal que } x - y = 3k$$

Demuestre que ~ es una relación de equivalencia y que el conjunto cociente $Z/\sim$ está formado por las clases de equivalencia [0], [1], [2].

27. Sea X el conjunto de todas las rectas en el plano. En X defina la relación ~ como

  a) $x \sim y$ si x es paralela a y

  b) $x \sim y$ si x es perpendicular a y

  Acepte que toda recta es paralela a sí misma. Demuestre entonces que la relación definida en a) es una relación de equivalencia, ¿cuáles son las clases de equivalencia así generadas? ¿Es la relación definida en b) una relación de equivalencia?

28. Sea X el conjunto de todos los triángulos. Defina en X la relación ~ como

  a) $x \sim y$ si x es semejante a y

  b) $x \sim y$ si x es congruente con y

  Demuestre que éstas son relaciones de equivalencia en X.

29. En el conjunto de números enteros **Z** se ha definido una cierta relación de equivalencia ~ de modo que **Z** ha quedado descompuesto como la unión de los siguientes conjuntos (disconjuntos)

$$\{\ldots, -10, -5, 0, 5, 10, \ldots\}$$
$$\{\ldots, -9, -4, 1, 6, 11, \ldots\}$$
$$\{\ldots, -8, -3, 2, 7, 12, \ldots\}$$
$$\{\ldots, -7, -2, 3, 8, 13, \ldots\}$$
$$\{\ldots, -6, -1, 4, 9, 14, \ldots\}$$

  Describa la relación ~.

30. En el conjunto N × N defina la relación ~ como

$$(a,b) \sim (c,d) \Leftrightarrow a + d = b + c$$

  Pruebe que ~ es una relación de equivalencia.

② 31. El siguiente argumento "demuestra" que si una relación binaria R es un conjunto X es simétrica y transitiva, entonces es reflexiva: sean $x, y \in X$ tales que xRy. Como R es simétrica y se tiene yRx. Como R es transitiva, xRy y yRx implica xRx, lo que muestra que R es reflexiva.

  Encuentre el error en este razonamiento.

---

32. Considere los conjuntos

$$A = \{a,b,c,d\}$$
$$B = \{1,2,3,4\}$$

  Se dan a continuación conjuntos de parejas ordenadas (x, y), $x \in A$, $y \in B$. Diga en cada caso si se tiene una función $f: A \rightarrow B$ en donde $y = f(x)$. En caso afirmativo determine

  1) el rango de la función

  2) si la función es inyectiva

  3) si la función tiene inversa. En tal caso determine la función $f^{-1}: B \rightarrow A$

  a) $\{(a,1), (b,2), (c,3), (d,4)\}$

  b) $\{(a,1), (a,2), (b,3), (b,4), (c,1), (c,2)\}$

  c) $\{(a,3), (b,3), (c,3), (d,3)\}$

  d) $\{(a,1), (b,4), (c,3), (d,2)\}$

  e) $\{(a,1), (b,3), (a,1), (c,2), (d,2)\}$

  f) $\{(d,1), (d,2), (a,2), (b,3), (c,4)\}$

  g) $\{(a,4), (d,3), (c,2), (b,1)\}$

  h) $\{(a,1)\ (b,3), (a,2), (c,3), (d,2)\}$

  i) $\{(b,3), (c,2), (a,2), (d,2)\}$

  j) $\{(c,1), (d,1), (a,2), (b,2)\}$

33. Sean las funciones $f, g: R \rightarrow R$ dadas por

$$f(x) = x^3 + x - 1$$
$$g(x) = a(x-1)^3 + b(x+2)^2 + cx + d$$

  Determine a, b, c y d para que $f = g$.

34. Se dice que la función $f: R \rightarrow R$ es *lineal* si $f(cx_1 + x_2) = cf(x_1) + f(x_2)$, en donde c_1, $x_1, x_2 \in R$. Determine cuáles de las siguientes funciones son lineales:

  a) $f(x) = x$

  b) $f(x) = 3x^2$

  c) $f(x) = x + 1$

  d) $f(x) = 2x^3 + 1$

  e) $f(x) = mx$, m una constante

  f) $f(x) = mx + b$, m y b constantes

35. Demuestre que la función $f: A \rightarrow B$ es una biyección si, y sólo si cumple con la siguiente propiedad: dado $b \in B$ el conjunto $f^{-1}(b) \subset A$ consta de un solo elemento.

① 36. Sea A un conjunto *finito* y considere una función $f: A \rightarrow A$. Demuestre que f es inyectiva si, y sólo si f es biyectiva.

37. Precise todas las posibles funciones biyectivas del conjunto $\{1, 2, 3\}$ en sí mismo (véase ejercicio 36).

38. Si A es un conjunto de n elementos, ¿cuántas funciones biyectivas de A en A distintas existen?

39. Sean $f: A \rightarrow B$ y $g: B \rightarrow C$ dos funciones. Demuestre que

  a) si f y g son inyectivas entonces $g \circ f$ es inyectiva.

  b) si f y g son sobreyectivas entonces $g \circ f$ es sobreyectiva.

  c) si f y g son biyectivas entonces $g \circ f$ es biyectiva.

① 40. Suponga que las funciones del ejercicio anterior son tales que $g \circ f$ es una función: a) inyectiva, b) sobreyectiva, c) biyectiva. ¿Qué puede decir de las funciones f y g?

41. Sea $f: A \rightarrow B$ una función biyectiva. Demuestre que la función $f^{-1}: B \rightarrow A$ tiene inversa y que $(f^{-1})^{-1} = f$.

42. Sean $f: A \rightarrow B$ y $g: B \rightarrow C$ dos funciones biyectivas, pruebe que la función $g \circ f: A \rightarrow C$ tiene inversa y que $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

43. Sean $f: A \rightarrow B$, $g: B \rightarrow C$, $h: C \rightarrow D$ tres funciones. Demuestre que las funciones $h \circ (g \circ f)$, $(h \circ g) \circ f: A \rightarrow D$ son iguales.

② 44. Considere la función: f: N → N tal que

$$f(x + y) = f(x) + f(y),\ \forall\, x, y \in \mathrm{N}$$
$$f(1) = 2$$

  Compruebe que $f(x) = 2x\ \forall\, x \in \mathrm{N}$.

① 45. Sea $f: A \to B$ una función de A a B y sean X y Y subconjuntos de A y B respectivamente. Demuestre que

a) $f(f^{-1}(Y)) \subseteq Y$

b) $f(f^{-1}(Y)) = Y$ si, y sólo si f es sobreyectiva.

c) $X \subseteq f^{-1}(f(X))$

d) $X = f^{-1}(f(X))$ si, y sólo si f es inyectiva.

① 46. Sea $f: A \to B$ una función de A a B y sean X_1 y X_2 subconjuntos de A. Demuestre que

a) $f(X_1 \cup X_2) = f(X_1) \cup f(X_2)$

b) $f(X_1 \cap X_2) \subseteq f(X_1) \cap f(X_2)$

Compruebe que $f(X_1 \cap X_2) = f(X_1) \cap f(X_2)$ si, y sólo si f es inyectiva.

CAPÍTULO UNO

Sistemas de ecuaciones lineales y matrices

En este capítulo se exponen las ideas y resultados preliminares que aparecen en el estudio del álgebra lineal. Se discutirá, por una parte, el método de eliminación de Gauss-Jordan para resolver sistemas de ecuaciones lineales, siendo éste una herramienta fundamental que se usará en el resto del libro. Por otra parte, se introducirá formalmente el concepto de matriz y se hará un estudio más o menos detallado de las principales propiedades que tienen estos "entes matemáticos" llamados matrices, que aparecen también como elementos fundamentales de trabajo en el álgebra lineal.

Para introducirse en el tema, se retoma la tan familiar situación de resolución de un sistema de dos ecuaciones lineales con dos incógnitas (x y y), dígase:

$$ax + by = m$$
$$cx + dy = n$$

(donde a, b, c, d, m, n son números reales dados).

Uno de los procedimientos más conocidos para obtener la solución de este sistema es el llamado "método de eliminación", que consiste en, por medio de operaciones algebraicas (permisibles), *eliminar* una de las incógnitas y producir así una nueva ecuación en la que aparece entonces sólo la otra. Por ejemplo, si se multiplica por $(-c)$ la primera ecuación y por (a) la segunda, y luego se suman, se obtiene

$$(ad - bc)y = an - cm$$

de donde se ve que si $ad - bc \neq 0$ entonces el sistema posee por solución

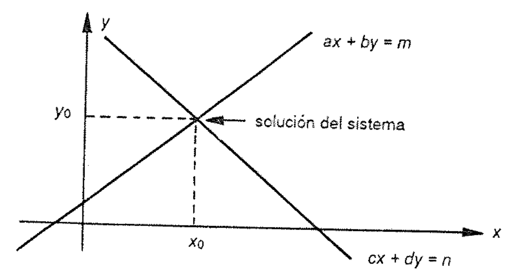
$$y = \frac{an - cm}{ad - bc}$$

$$x = \frac{dm - bn}{ad - bc}$$

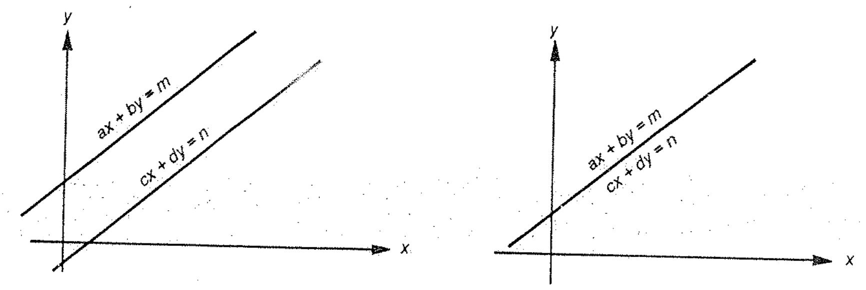
(el valor de x se obtiene sustituyendo el valor para y encontrado en cualquiera de las ecuaciones y despejando).

Desde el punto de vista geométrico, cada una de las ecuaciones del sistema dado representa una recta en el plano. Hallar una solución del sistema consiste en localizar el punto en el que tales rectas se intersecan (punto cuyas coordenadas

satisfacen simultáneamente ambas ecuaciones). Así pues, si $ad - bc \neq 0$ se tiene la siguiente situación:



mientras que si $ad - bc = 0$, puede acontecer que las dos rectas sean paralelas (en cuyo caso se dice que el sistema no tiene solución) o bien que se trate de la misma recta (en cuyo caso se dice que el sistema posee una infinidad de soluciones). Gráficamente se ve como



La primera parte de este capítulo (secciones 1 y 2) no es más que una generalización de las ideas involucradas en este sencillo ejemplo. Por ejemplo, se tratará de dar un sentido más amplio (¡y preciso!) a los conceptos: ecuación lineal, sistema de ecuaciones lineales, solución de un sistema de ecuaciones lineales, etc., que son términos que ya se usan en este ejemplo elemental, así como también se establecerá un método general para resolver sistemas de ecuaciones. En la segunda parte de este capítulo, se introduce formalmente el concepto de "matriz" y se estudian las propiedades más importantes de ellas (secciones 3 y 4).

1. SISTEMAS DE ECUACIONES LINEALES

Se llama *ecuación lineal* en las incógnitas (o indeterminadas) $x_1, x_2, \ldots, x_n$* a una ecuación de la forma

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b \tag{1.1}$$

*En el caso de una ecuación lineal con dos incógnitas se usan tradicionalmente las letras x y y para denotar a éstas, así como en el caso de 3, 4 o 5 incógnitas se usan también, además de x y y, las letras z, u, v, w. En el caso general se usará (por razones obvias) la misma letra x con un subíndice que indicará el número de incógnita de la que se trate. Así, la ecuación $ax + by = m$ se verá como $ax_1 + bx_2 = m$, y la ecuación $ax + by + cz + du + ew = n$ se verá como $ax_1 + bx_2 + cx_3 + dx_4 + ex_5 = n$.

en donde $a_1, a_2, \ldots, a_n, b$ son números (reales) dados.

Al número a_i que multiplica a la incógnita x_i $(i = 1, 2, \ldots, n)$ se le llama *coeficiente* de la incógnita correspondiente. Al número b se le llama *término independiente* de la ecuación.

Si $b = 0$, se dice que la ecuación (1.1) es *homogénea*. En caso contrario, se dice que es *no homogénea*.

EJEMPLO 1

Así pues, las ecuaciones $3x_1 + 2x_2 = 6$, $5x_1 + 6x_2 = 0$ son ecuaciones lineales en las incógnitas x_1, x_2 (la primera de ellas no homogénea y la segunda homogénea), mientras que la ecuación $3x_1^2 + 2x_2 = 1$ no es una ecuación lineal.

Se dice que $x_1 = \tilde{x}_1, x_2 = \tilde{x}_2, \ldots, x_n = \tilde{x}_n$ $(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ son números determinados) es una *solución* de la ecuación (1), si estos valores de las incógnitas *satisfacen* la ecuación, en el sentido de que sustituyendo estos valores en ella, ésta se reduce a una identidad, esto es, si

$$a_1\tilde{x}_1 + a_2\tilde{x}_2 + \ldots + a_n\tilde{x}_n \equiv b$$

EJEMPLO 2

Por ejemplo, para la ecuación $3x_1 = 6$, se tiene que $x_1 = 2$ es una solución, puesto que $3(2) = 6$ es una identidad, así como para la ecuación $x_1 + 3x_2 = 4$, se tiene que $x_1 = x_2 = 1$ y $x_1 = 10$, $x_2 = -2$ son soluciones, ya que $1 + 3(1) = 4$ y $10 + 3(-2) = 4$ son identidades.

Se llamará *conjunto solución* de la ecuación (1.1) al conjunto formado por las soluciones de la ecuación.

EJEMPLO 3

Por ejemplo, para una ecuación lineal con una incógnita x, $ax = b$ se pueden tener las siguientes situaciones:

a) Si $a = 2$, $b = 4$ (la ecuación es pues $2x = 4$) el conjunto solución está formado por el único valor de la incógnita $x = 2$.

b) Si $a = 0$, $b = 0$ (la ecuación es pues $0 \cdot x = 0$) el conjunto solución está formado por una infinidad de valores de x; esto es, cualquier valor de x satisface la ecuación.

c) Si $a = 0$, $b = 1$ (la ecuación es pues $0 \cdot x = 1$) el conjunto solución es vacío, pues no existe valor $\tilde{x}$ de x que cumpla con $0 \cdot \tilde{x} = 1$.

Véase pues que en este ejemplo, el conjunto solución es vacío, o contiene un único valor de x, o está formado por una infinidad de valores para x (obsérvese que no puede ocurrir que el conjunto solución tenga dos o tres o n -n finito mayor que 1- valores de x). En la sección 3 de este capítulo se verá que éste es un hecho general para sistemas de ecuaciones lineales.

En esta sección y en la siguiente, será de especial interés considerar *sistemas* de ecuaciones lineales.

Al conjunto de m ecuaciones lineales

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \ldots \end{aligned} \tag{1.2}$$

$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m$$

en donde a_{ij}, $b_i(1 \leq i \leq m,\ 1 \leq j \leq n)$ son números dados, se le llama *sistema de m ecuaciones lineales en las n incógnitas* $x_1, \ldots x_n$. Si los términos independientes $b_i(i = 1, 2, \ldots n)$ son todos iguales a cero, se dice que el sistema es *homogéneo.* En caso contrario se dice que es *no homogéneo*.

Obsérvese que se tiene que introducir un nuevo subíndice a los coeficientes de las ecuaciones: el primero de ellos indica el número de ecuación de la que se trata y el segundo, el número de la incógnita de la que es coeficiente. Así pues, a_{ij} es el coeficiente de la incógnita x_j en la i-ésima ecuación.

Los conceptos de solución y de conjunto solución para el sistema (1.2) se generalizan de una manera natural: a los valores $x_i = \tilde{x}_i\ (i = 1, \ldots, n)$ se les llama solución del sistema si estos satisfacen todas y cada una de las ecuaciones del mismo (o sea, si son soluciones para todas y cada una de las ecuaciones). El conjunto solución está constituido por las soluciones del sistema.

EJEMPLO 4 Por ejemplo, el sistema

$$3x_1 + 2x_2 = 7$$

$$x_1 + x_2 = 3$$

tiene por conjunto solución a $\{x_1 = 1, x_2 = 2\}$ (la solución es "única").

EJEMPLO 5

$$x_1 + x_2 = 3$$

$$2x_1 + 2x_2 = 6$$

El sistema posee una infinidad de soluciones, pues dando cualquier valor a x_1, dígase $x_1 = t\ (t \in R)$, basta con escribir $x_2 = 3 - t$ para que se satisfaga el sistema. En este caso el conjunto solución es $\{x_1 = t, x_2 = 3 - t, t \in R\}$ (el sistema tiene una infinidad de soluciones).

EJEMPLO 6 Por último, es claro que el sistema

$$x_1 + 2x_2 = 1$$

$$x_1 + 2x_2 = 3$$

no posee soluciones, esto es, el conjunto solución es vacío.

Una vez dados estos conceptos preliminares, se pasa a la discusión de *cómo* obtener las soluciones (si existen) de un sistema como (1.2). En el caso $m = n = 2$ (sistema de dos ecuaciones con dos incógnitas) se sabe perfectamente qué hacer: se realizan operaciones algebraicas en las ecuaciones con el objeto de eliminar una de las incógnitas y se resuelve para la restante. Para el caso general, se verá que "la filosofía" del procedimiento es esencialmente la misma: tratar de eliminar incógnitas.

En una primera vista, parece pues que no hay nada nuevo qué decir: dado un sistema de ecuaciones lineales, se realizan operaciones en las ecuaciones de tal manera que se eliminan incógnitas ¡y ya! Resulta que, efectivamente, no hay nada nuevo qué decir, *a excepción* de la puesta en escena de tal estrategia.

Se debe ser muy cuidadoso con el tipo de operaciones que se realizan en las ecuaciones: considérese por ejemplo el sistema de 3 ecuaciones con 2 incógnitas

$$5x - 2y = 1$$

$$-x + 2y = 2$$

$$4x + 2y = 9$$

y aplíquese libremente la "estrategia de eliminar incógnitas". Sumando la segunda y tercera ecuación se tiene

$$3x + 4y = 11$$

que junto con la primera de ellas $5x - 2y = 1$, produce (dígase que se multiplica por 2 esta última y se suma con $3x + 4y = 11$ eliminando así a y) por "solución" $x = 1, y = 2$.

Si se verifica esta "solución" del sistema en las ecuaciones, se ve que, a excepción de la primera de ellas, para la cual se tiene $5(1) - 2(2) = 1$, una identidad, en las dos restantes estos valores de x y y *no* conducen a identidades (se tiene $-(1) + 2(2) \neq 2$, $4(1) + 2(2) \neq 9$), y por tanto, lo que se obtiene *no fue una solución del sistema.*

Ya en la introducción al capítulo se observa que en el método de eliminación se deben realizar operaciones algebraicas *permisibles* en las ecuaciones para eliminar incógnitas de ellas. Uno de los puntos que se discutirán aquí es acerca de *cuáles* son esas operaciones permisibles (se observará después, que en el ejemplo anterior se realiza una operación no permisible en las ecuaciones del sistema). También se discutirá la metodología para realizar tales operaciones.

Para concretar la discusión, se tomará como ejemplo un sistema de 3 ecuaciones con 3 incógnitas.

Si se observa el sistema

$$\begin{aligned} x_1 &= 1 \\ x_2 &= -1 \\ x_3 &= 4 \end{aligned} \tag{1.3}$$

y se pregunta por la solución de éste, se ve que no se tiene nada que hacer, pues la solución está escrita en el mismo sistema. ¿Qué es lo que tienen de particular las ecuaciones de este sistema? Sucede que en cada una de ellas aparece sólo una incógnita y ésta con coeficiente 1 (esto es, en cada ecuación aparece cada una de las indeterminadas *despejada*).

Sin embargo, si se observa el sistema

$$\begin{aligned} 2x_1 + 3x_2 - 5x_3 &= -21 \\ x_1 + x_2 + 3x_3 &= 12 \\ 5x_1 + 2x_2 + 5x_3 &= 23 \end{aligned} \tag{1.4}$$

no se puede, a simple vista, decir cuáles son sus soluciones.

El método de eliminación permitirá pasar del sistema (1.4) al sistema (1.3); [se verifica fácilmente que (1.3) es solución de (1.4)].

1.1. SISTEMAS EQUIVALENTES Y EL MÉTODO DE ELIMINACIÓN

DEFINICIÓN 1.1 Se dice que dos sistemas de ecuaciones son *equivalentes* si ambos poseen el mismo número de ecuaciones e incógnitas y poseen también el mismo conjunto solución.

Por ejemplo, el sistema (1.3) y el sistema (1.4) son equivalentes.

Dado entonces un sistema de ecuaciones por resolver, se deben realizar operaciones algebraicas en las ecuaciones persiguiendo 2 fines:

1) Obtener un sistema más simple que el dado [en el sentido en el que (1.3) es más simple que (1.4)].
2) Que el sistema así obtenido sea equivalente al original.

Ésta es la filosofía propiamente dicha del método de eliminación para resolver sistemas de ecuaciones.

De esta estrategia general surge una interrogante natural:

¿Qué tipo de operaciones se pueden realizar en las ecuaciones para producir sistemas equivalentes?

Véase que son 3 operaciones, llamadas *operaciones elementales*, que permitirán obtener sistemas equivalentes. Éstas son:

1) Intercambiar de posición 2 de las ecuaciones.
2) Multiplicar una de las ecuaciones por una constante diferente de cero.
3) Sumar a una ecuación un múltiplo de otra.

En efecto:

Es claro que si en un sistema de ecuaciones se intercambia la posición de dos de ellas, el nuevo sistema tendrá las mismas soluciones que el inicial. Así, en el sistema (1.4), es claro que, por ejemplo intercambiando la primera y segunda ecuación, se obtiene el sistema

$$\begin{aligned} x_1 + x_2 + 3x_3 &= 12 \\ 2x_1 + 3x_2 - 5x_3 &= -21 \\ 5x_1 + 2x_2 + 5x_3 &= 23 \end{aligned} \qquad (1.4.a)$$

el cual posee el mismo conjunto solución que (1.4).

Por otra parte, véase que al multiplicar una de las ecuaciones de un sistema por una constante diferente de cero, se obtiene un nuevo sistema equivalente. Piénsese en general en el sistema (1.2) y supóngase que a la i-ésima ecuación se le ha multiplicado por la constante $k \neq 0$, quedando entonces el sistema

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$

$$ka_{i1}x_1 + ka_{i2}x_2 + \ldots + ka_{in}x_n = kb_i \qquad (1.2.a)$$

$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m$$

Para verificar que (1.2) y (1.2.a) son sistemas equivalentes, sólo se debe tener cuidado con la i-ésima ecuación que es la que ha sido alterada. Supóngase que $x_j = \tilde{x}_j$ ($j = 1, 2, \ldots, n$) es solución de (1.2). Entonces

$$a_{i1}\tilde{x}_1 + a_{i2}\tilde{x}_2 + \ldots + a_{in}\tilde{x}_n = b_i \qquad (1.5)$$

es una identidad.

Si se multiplica esta identidad por k se obtiene

$$ka_{i1}\tilde{x}_1 + ka_{i2}\tilde{x}_2 + \ldots + ka_{in}\tilde{x}_n = kb_i \qquad (1.6)$$

la cual sigue siendo una identidad, lo que dice entonces que $x_j = \tilde{x}_j$ ($j = 1, 2, \ldots, n$) es también solución de (1.2.a). Por último, como la constante k es distinta de cero, se puede, a partir de (1.6), regresar a (1.5) multiplicando por k^{-1}, y entonces se ve que cualquier solución de (1.2.a) es también solución de (1.2). En resumen, los sistemas (1.2) y (1.2.a) son sistemas equivalentes.

Por último, para ver que sumando un múltiplo de una ecuación a otra no se altera el conjunto solución, supóngase, sin pérdida de generalidad, que en el sistema (1.2) se suma k veces la primera ecuación a la segunda (más concretamente: se sustituye la segunda ecuación por ella misma más k veces la primera), quedando entonces esta segunda ecuación como

$$(a_{21} + ka_{11})x_1 + (a_{22} + ka_{12})x_2 + \ldots + (a_{2n} + ka_{1n})x_n = b_2 + kb_1 \qquad (1.7)$$

(todas las demás ecuaciones del sistema no se han alterado).

Supóngase que $x_j = \tilde{x}_j$ ($j = 1, 2, \ldots, n$) es solución del sistema (1.2). Entonces

$$a_{11}\tilde{x}_1 + a_{12}\tilde{x}_2 + \ldots + a_{1n}\tilde{x}_n = b_1$$

$$a_{21}\tilde{x}_1 + a_{22}\tilde{x}_2 + \ldots + a_{2n}\tilde{x}_n = b_2$$

son identidades. Si se suma k veces la primera de estas identidades a la segunda se obtiene

$$(a_{21} + ka_{11})\tilde{x}_1 + (a_{22} + ka_{12})\tilde{x}_2 + \ldots + (a_{2n} + ka_{1n})\tilde{x}_n = b_2 + kb_1$$

la cual es también una identidad. Esto quiere decir entonces que $x_j = \tilde{x}_j$ ($j = 1, 2, \ldots, n$) es también solución del nuevo sistema. Por último, obsérvese que se puede pasar del nuevo sistema al sistema original sumando (en el nuevo sistema) $-k$ veces la primera ecuación a la segunda (la ecuación (1.7), recuperando así la ecuación original), de modo que cualquier solución del nuevo sistema es también solución del sistema original. En conclusión, ambos sistemas son equivalentes.

EJEMPLO 7 Observe cómo funcionan estas ideas en nuestro ejemplo, el sistema (1.4). Se tratará entonces de realizar operaciones elementales en las ecuaciones (creando así sistemas equivalentes) que conduzcan a una simplificación en las mismas (intentando siempre eliminar incógnitas).

Parta entonces del sistema (1.4.a) proveniente del sistema (1.4) y equivalente al mismo

$$\begin{aligned} x_1 + x_2 + 3x_3 &= 12 \\ 2x_1 + 3x_2 - 5x_3 &= -21 \\ 5x_1 + 2x_2 + 5x_3 &= 23 \end{aligned}$$

Sustituya ahora la segunda ecuación por ella misma menos 2 veces la primera y la tercera por ella misma menos 5 veces la primera (logrando así eliminar la incógnita x_1 de estas ecuaciones). Queda entonces

$$\begin{aligned} x_1 + x_2 + 3x_3 &= 12 \\ x_2 - 11x_3 &= -45 \\ -3x_2 - 10x_3 &= -37 \end{aligned}$$

Sustituya ahora la tercera ecuación por ella misma más 3 veces la segunda, quedando ahora

$$\begin{aligned} x_1 + x_2 + 3x_3 &= 12 \\ x_2 - 11x_3 &= -45 \\ -43x_3 &= -172 \end{aligned}$$

Multiplique la tercera ecuación por $-\frac{1}{43}$

$$\begin{aligned} x_1 + x_2 + 3x_3 &= 12 \\ x_2 - 11x_3 &= -45 \\ x_3 &= 4 \end{aligned}$$

Sustituya la primera ecuación por ella misma menos 3 veces la tercera y la segunda por ella misma más 11 veces la tercera y obtenemos

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_2 &= -1 \\ x_3 &= 4 \end{aligned}$$

Finalmente sustituya la primera ecuación por ella misma menos la segunda y obtenga

$$\begin{aligned} x_1 &= 1 \\ x_2 &= -1 \\ x_3 &= 4 \end{aligned}$$

y así se ha logrado el sistema (1.3), el cual se sabe que tiene las mismas soluciones que el original (1.4). Por tanto, las soluciones (1.4) son $x_1 = 1$, $x_2 = -1$ y $x_3 = 4$.

EJEMPLO 8 Véase un ejemplo más. Se quiere resolver el sistema

$$\begin{aligned} 3x_1 - 2x_2 + 2x_3 + x_4 &= 1 && (E1) \\ x_1 - x_2 + x_3 - 3x_4 &= 0 && (E2) \\ 2x_1 + x_2 + 2x_3 + 4x_4 &= 5 && (E3) \\ 4x_1 - 4x_2 + x_3 + 2x_4 &= 4 && (E4) \end{aligned}$$

Para abreviar, se escribirá $(E_i) \leftrightarrow (E_j)$ para indicar que se ha intercambiado la posición de las ecuaciones (E_i) y (E_j), $(E_i) \rightarrow k(E_i)$ para indicar que se ha multiplicando por k la ecuación (E_i) y $(E_i) \rightarrow (E_i) + k(E_j)$ para indicar que se ha sustituido la ecuación (E_i) por ella misma más k veces la ecuación (E_j) (refiriéndose siempre al sistema inmediato anterior).

Procédase entonces a realizar operaciones elementales con el objeto de eliminar incógnitas y obtener un sistema equivalente del cual se puedan leer fácilmente las soluciones. Se irán escribiendo en la izquierda los sistemas obtenidos y en la derecha las operaciones elementales que se han realizado para obtenerlos

$$\begin{aligned} x_1 - x_2 - x_3 - 3x_4 &= 0 && (E1) \leftrightarrow (E2) \\ 3x_1 - 2x_2 + 2x_3 + x_4 &= 1 \\ 2x_1 + x_2 + 2x_3 + 4x_4 &= 5 \\ 4x_1 - 4x_2 + x_3 + 2x_4 &= 4 \end{aligned}$$

$$\begin{aligned} x_1 - x_2 - x_3 - 3x_4 &= 0 && (E2) \rightarrow (E2) - 3(E1) \\ x_2 + 5x_3 + 10x_4 &= 1 && (E3) \rightarrow (E3) - 2(E1) \\ 3x_2 + 4x_3 + 10x_4 &= 5 && (E4) \rightarrow (E4) - 4(E1) \\ 5x_3 + 14x_4 &= 4 \end{aligned}$$

$$\begin{aligned} x_1 - x_2 - x_3 - 3x_4 &= 0 && (E3) \rightarrow (E3) - 3(E2) \\ x_2 + 5x_3 + 10x_4 &= 1 \\ -11x_3 - 20x_4 &= 2 \\ 5x_3 + 14x_4 &= 4 \end{aligned}$$

$$\begin{aligned} x_1 - x_2 - x_3 - 3x_4 &= 0 && (E3) \rightarrow -\tfrac{1}{11}(E3) \\ x_2 + 5x_3 + 10x_4 &= 1 \\ x_3 + \tfrac{20}{11}x_4 &= -\tfrac{2}{11} \\ 5x_3 + 14x_4 &= 4 \end{aligned}$$

$$\begin{aligned} x_1 - x_2 - x_3 - 3x_4 &= 0 \\ x_2 + 5x_3 + 10x_4 &= 1 \\ x_3 + \tfrac{20}{11}x_4 &= -\tfrac{2}{11} \\ \tfrac{54}{11}x_4 &= \tfrac{54}{11} \end{aligned} \qquad (E4) \to (E4) - 5(E3)$$

$$\begin{aligned} x_1 - x_2 - x_3 - 3x_4 &= 0 \\ x_2 + 5x_3 + 10x_4 &= 1 \\ x_3 + \tfrac{20}{11}x_4 &= -\tfrac{2}{11} \\ x_4 &= 1 \end{aligned} \qquad (E4) \to \tfrac{11}{54}(E4)$$

$$\begin{aligned} x_1 - x_2 - x_3 &= 3 \\ x_2 + 5x_3 &= -9 \\ x_3 &= -2 \\ x_4 &= 1 \end{aligned} \qquad \begin{aligned} &(E1) \to (E1) + 3(E4) \\ &(E2) \to (E2) - 10(E4) \\ &(E3) \to (E3) - \tfrac{20}{11}(E4) \end{aligned}$$

$$\begin{aligned} x_1 - x_2 &= 1 \\ x_2 &= 1 \\ x_3 &= -2 \\ x_4 &= 1 \end{aligned} \qquad \begin{aligned} &(E1) \to (E1) + (E3) \\ &(E2) \to (E2) - 5(E3) \end{aligned}$$

$$\begin{aligned} x_1 &= 2 \\ x_2 &= 1 \\ x_3 &= -2 \\ x_4 &= 1 \end{aligned} \qquad (E1) \to (E1) + (E2)$$

La solución del sistema es pues $x_1 = 2, x_2 = 1, x_3 = -2, x_4 = 1$.

EJERCICIOS (SECCIÓN 1, CAPÍTULO 1)

1. Determine cuáles de las siguientes ecuaciones son lineales:
   - a) $3x + 2y = z$
   - b) $5x_1 + 2x_2 + 3x_3 = 0$
   - c) $6x_1 = 3x_2$
   - d) $3x^2 + 2y = 5$
   - e) $5x_1 + 6x_2 + 7x_3^3 = 7$
   - f) $6x_1 - 7x_2 + 8x_3 = 4$

2. En los siguientes incisos se da una ecuación lineal y un conjunto de valores de las indeterminadas de la ecuación.
   Determine si los elementos del conjunto dado son soluciones de la ecuación correspondiente.
   - a) $5x = 7$ $\quad \{x = 2\}$
   - b) $3x + 6y = 4$ $\quad \{x = 2, y = 1\}$
   - c) $5x + 2y = 0$ $\quad \{x = 0, y = 0\}$
   - d) $2x + 4y = 0$ $\quad \{x = 1, y = -\frac{1}{2}\}$
   - e) $x + 3y = 0$ $\quad \{x = t, y = -\frac{1}{3}t, t \in \mathbf{R}\}$
   - f) $4x + 3y - z = 4$ $\quad \{x = t, y = s, z = 4t + 3s - 4, t, s \in \mathbf{R}\}$
3. Determine el conjunto solución de la ecuación $0 \cdot x = 4$.
4. Determine el conjunto solución de la ecuación $5x = 10$.
5. Determine el conjunto solución de la ecuación $ax + by = c$ en cada uno de los casos:
   - a) $a = 0, b = 0, c = 0$
   - b) $a = 0, b = 0, c \neq 0$
   - c) $a = 0, b \neq 0, c = 0$
   - d) $a = 0, b \neq 0, c \neq 0$
   - e) $a \neq 0, b = 0, c = 0$
   - f) $a \neq 0, b = 0, c \neq 0$
   - g) $a \neq 0, b \neq 0, c = 0$
   - e) $a \neq 0, b \neq 0, c \neq 0$
6. Determine el conjunto solución de los siguientes sistemas de dos ecuaciones con dos incógnitas. En cada caso, interprete geométricamente las ecuaciones del sistema, así como su solución.
   - a) $\begin{aligned} 3x + 2y &= 0 \\ x - y &= 4 \end{aligned}$
   - b) $\begin{aligned} 5x + y &= 1 \\ 5x + y &= 0 \end{aligned}$
   - c) $\begin{aligned} 3x + 2y &= 4 \\ 6x + 4y &= 8 \end{aligned}$
   - d) $\begin{aligned} x + y &= 0 \\ x - y &= 0 \end{aligned}$
   - e) $\begin{aligned} x + y &= 0 \\ 5x + 5y &= 0 \end{aligned}$
   - f) $\begin{aligned} 3x + 5y &= 8 \\ 5x - 2y &= 3 \end{aligned}$
7. Verifique que el conjunto dado está formado por soluciones del sistema correspondiente.
   - a) $\begin{aligned} 3x_1 + 2x_2 - 5x_3 &= 0 \\ x_1 + x_2 - x_3 &= 1 \\ 2x_1 - x_2 + 6x_3 &= 7 \end{aligned}$ $\quad \{x_1 = x_2 = x_3 = 1\}$
   - b) $\begin{aligned} x_1 + 2x_2 + 6x_3 - 4x_4 &= 4 \\ 2x_1 + 7x_2 - 4x_3 + 9x_4 &= 9 \\ x_1 + 3x_2 + x_3 - 6x_4 &= -3 \\ 6x_1 - 4x_2 + 4x_3 - 10x_4 &= 6 \end{aligned}$ $\quad \{x_1 = 2, x_2 = 0, x_3 = x_4 = 1\}$
   - c) $\begin{aligned} x_1 - x_2 + x_3 &= 1 \\ 2x_1 - 2x_2 + 2x_3 &= 2 \end{aligned}$ $\quad \{x_1 = 1 + t - s, x_2 = t, x_3 = s, t, s \in \mathbf{R}\}$
8. Considere los siguientes sistemas de ecuaciones lineales:

$$\begin{aligned} 2x_1 + 3x_2 &= 5 \\ x_1 - 4x_2 &= -3 \end{aligned} \qquad \text{y} \qquad \begin{aligned} 2x_1 + 3x_2 &= 5 \\ x_1 - 4x_2 &= -3 \\ 2x_1 - 8x_2 &= -6 \end{aligned}$$

Determine el conjunto solución de cada uno de estos sistemas. ¿Son sistemas equivalentes?

9. Considere el sistema de 3 ecuaciones con 2 incógnitas.

$$\text{(A)}\qquad \begin{aligned} 5x - 2y &= 1 \\ -x + 2y &= 2 \\ 4x + 2y &= 9 \end{aligned}$$

Al sustituir la segunda y tercera ecuación por la suma de ellas se obtiene el sistema

$$\text{(B)}\qquad \begin{aligned} 5x - 2y &= 1 \\ 3x + 4y &= 11 \end{aligned}$$

Demuestre que toda solución del sistema (A) es también solución del sistema (B). Es decir, compruebe que el conjunto solución del sistema (A) es un subconjunto del conjunto solución del sistema (B). ¿Es válida también la otra contención? ¿Son (A) y (B) sistemas equivalentes? Explique.

10. Pruebe que los sistemas de ecuaciones

$$\begin{aligned} 5x_1 - 4x_2 &= 2 \\ 2x_1 + 3x_2 &= 10 \end{aligned} \qquad \text{y} \qquad \begin{aligned} x_1 &= 2 \\ x_2 &= 2 \end{aligned}$$

son equivalentes.

11. Compruebe que el sistema

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

es equivalente al sistema

$$\begin{aligned} (k_1a_{11} + k_2a_{21})x_1 + (k_1a_{12} + k_2a_{22})x_2 &= k_1b_1 + k_2b_2 \\ (m_1a_{11} + m_2a_{21})x_1 + (m_1a_{12} + m_2a_{22})x_2 &= m_1b_1 + m_2b_2 \end{aligned}$$

en donde k_1, k_2, m_1 y m_2 son constantes no nulas.

12. Use el método de eliminación discutido en esta sección para obtener la solución de los siguientes sistemas de ecuaciones:

a) $\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ 2x_1 - 4x_2 + 2x_3 &= 0 \\ 3x_1 - 2x_2 + 5x_3 &= 14 \end{aligned}$

b) $\begin{aligned} x_1 + x_2 + 3x_3 &= 3 \\ 2x_1 + 5x_2 + x_3 &= -2 \\ -x_1 + x_2 + 4x_3 &= 6 \end{aligned}$

c) $\begin{aligned} 2x_1 - 2x_2 + x_3 + x_4 &= 4 \\ x_1 + x_2 + x_3 + 2x_4 &= 8 \\ 3x_1 - x_2 + x_3 - 2x_4 &= 0 \\ x_1 + 4x_2 + 3x_3 - 4x_2 &= 3 \end{aligned}$

2. EL MÉTODO DE ELIMINACIÓN DE GAUSS-JORDAN

El objetivo de esta sección es exponer el método de eliminación de Gauss-Jordan para resolver sistemas de ecuaciones lineales.

El método no es más que una "optimización y sistematización" de las ideas expuestas en la sección anterior.

El primer paso que se dará para lograr este objetivo va en la dirección de "ahorro de notación". Para aclarar esta idea, se debe recordar lo que motivó el estudio de la "división sintética" en los cursos de álgebra elemental.

Supóngase que se quiere dividir el polinomio $3x^5 - 2x^4 + 4x^3 - 3x^2 + 2x - 1$ entre el polinomio $x - 1$. Realícense entonces las operaciones estándar del algoritmo de la división de polinomios

$$\begin{array}{r|l} & 3x^4 + x^3 + 5x^2 + 2x + 4 \quad \text{(cociente)} \\ \hline x - 1 & 3x^5 - 2x^4 + 4x^3 - 3x^2 + 2x - 1 \\ & \underline{-(3x^5 - 3x^4)} \\ & \quad x^4 + 4x^3 \\ & \quad \underline{-(x^4 - x^3)} \\ & \qquad 5x^3 - 3x^2 \\ & \qquad \underline{-(5x^3 - 5x^2)} \\ & \qquad\quad 2x^2 + 2x \\ & \qquad\quad \underline{-(2x^2 - 2x)} \\ & \qquad\qquad 4x - 1 \\ & \qquad\qquad \underline{-(4x - 4)} \\ & \qquad\qquad\quad 3 \quad \text{(resto)} \end{array}$$

y se escribe entonces el resultado como

$$\frac{3x^5 - 2x^4 + 4x^3 - 3x^2 + 2x - 1}{x - 1} = 3x^4 + x^3 + 5x^2 + 2x + 4 + \frac{3}{x - 1}$$

Después de hacer repetidas veces este proceso, se da uno cuenta que existe una forma de hacer la misma división con un ahorro notable en la notación: toda la información en el proceso de la división está contenida en los coeficientes del polinomio $3x^5 - 2x^4 + 4x^3 - 3x^2 + 2x - 1$ y del polinomio $x - 1$.

Se llega así a la división sintética: la misma operación de división de estos polinomios puede efectuarse *haciendo las operaciones sólo con los coeficientes*. El algoritmo se verá ahora*

Coeficientes del numerador ↘

$$\begin{array}{cccccc|l} 3 & -2 & 4 & -3 & 2 & -1 & \\ \downarrow & 3 & 1 & 5 & 2 & 4 & 1 \\ \hline 3 & 1 & 5 & 2 & 4 & 3 & \end{array}$$

Valor de a en $x - a$

Coeficientes del cociente (3 1 5 2 4) — resto (3)

y se ve entonces que se llega al mismo resultado evitando escribir todas las x, lo cual representa un considerable ahorro de espacio y de tinta.

Al resolver un sistema de ecuaciones ocurre una situación similar: toda la información en el proceso de eliminación de incógnitas expuesto en la sección

*El algoritmo es

$$\begin{array}{ccccc|l} a_1 & a_2 & a_3 & \dots & a_n & b \\ & a_1b & (a_2+a_1b)b & & & \\ \hline a_1 & a_2+a_1b & a_3+(a_2+a_1b)b \dots & & & \end{array}$$

anterior, está contenida en los coeficientes de las ecuaciones. De nuevo, se puede evitar escribir las incógnitas $x_1, x_2, \ldots x_n$ y trabajar sólo con los coeficientes de las ecuaciones. Surge, de este modo, el concepto de matriz.*

2.1. MATRICES Y SISTEMAS DE ECUACIONES

Se llama *matriz* a un arreglo rectangular de números como el siguiente:

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \tag{2.1}$$

A los números a_{ij} ($i = 1, 2, \ldots, m, j = 1, 2, \ldots, n$) se les llama *elementos* de la matriz.

Se denota a las matrices con letras mayúsculas y a sus elementos con minúsculas.

Se escribe entonces la matriz A como

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \tag{2.2}$$

Otra notación que se usará frecuentemente para matrices es:

$A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$, la cual representa lo mismo que (2.2).

Se dirá que los elementos $a_{i1}, a_{i2} \ldots a_{in}$ constituyen la i-ésima *línea* de la matriz A, mientras que los elementos $a_{1j}, a_{2j}, \ldots, a_{mj}$ constituyen la j-ésima *columna* de la matriz A.

Así pues, el elemento a_{ij} se encuentra en la i-ésima línea y en la j-ésima columna de A

$$A = \begin{bmatrix} & a_{1j} & \\ & \vdots & \\ a_{i1} \ldots & a_{ij} & \ldots a_{in} \\ & \vdots & \\ & a_{mj} & \end{bmatrix} \quad i\text{-ésima línea}$$

j-ésima columna

*Aunque el concepto de matriz aparece en este momento sólo por una cuestión de ahorro en la notación en el proceso de solución de un sistema de ecuaciones, véase más adelante que en realidad estos elementos matemáticos tienen "vida propia" la cual se estudiará detenidamente.

El *orden* ("tamaño") de una matriz A se define como

número de líneas × número de columnas

Así, la matriz (2.2) tiene orden $m \times n$.

Las matrices A y B se dicen iguales si: 1) tienen el mismo orden, y 2) sus elementos correspondientes son iguales (o sea, si

$A = (a_{ij})_{\substack{i=1,\ldots,m_1 \\ j=1,\ldots,n_1}}$, $B = (b_{ij})_{\substack{i=1,\ldots,m_2 \\ j=1,\ldots,n_2}}$, entonces $A = B$ si $m_1 = m_2, n_1 = n_2$ y $a_{ij} = b_{ij}$, $i = 1, \ldots, m_1, j = 1, \ldots, n_1$).

Dado el sistema de m ecuaciones con n incógnitas

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned} \tag{2.3}$$

Se asocia a éste la matriz A (de orden $m \times n$) de sus coeficientes

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \tag{2.4}$$

llamada *matriz de coeficientes* (o matriz del sistema), así como la matriz $m \times (n + 1)$.

$$A_a = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ & & \ldots & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right] \tag{2.5}$$

(la misma matriz A con una última columna añadida, la de los términos independientes), llamada *matriz aumentada de coeficientes* (o del sistema).

EJEMPLO 1 Así pues, para el sistema de 2 ecuaciones con 4 incógnitas

$$\begin{aligned} 2x_1 - x_2 + 5x_3 - 3x_4 &= 1 \\ x_1 - 5x_2 + 10x_3 - 2x_4 &= -7 \end{aligned}$$

Se asocia la matriz del sistema(2 × 4)

$$A = \begin{bmatrix} 2 & -1 & 5 & -3 \\ 1 & -5 & 10 & -2 \end{bmatrix}$$

y la matriz aumentada (2 × 5)

$$A_a = \left[\begin{array}{cccc|c} 2 & -1 & 5 & -3 & 1 \\ 1 & -5 & 10 & -2 & -7 \end{array}\right]$$

EJEMPLO 2 Del mismo modo, para el sistema de 3 ecuaciones con 3 incógnitas

$$x_1 = 2$$
$$x_2 = 1$$
$$x_3 = 4$$

la matriz del sistema es:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

y la matriz aumentada

$$A_a = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

Antes de comenzar a describir la metodología de la eliminación Gaussiana, se introducirá el concepto de matriz en forma *escalonada reducida*.

Se dice que la matriz A está en la forma *escalonada reducida* si en ella se cumplen las siguientes 4 condiciones:

1) Si la matriz posee alguna o algunas líneas que consten exclusivamente de ceros, éstas se encuentran concentradas en la parte inferior de la matriz.
2) Para cada línea de la matriz que no consta exclusivamente de ceros, el primer elemento distinto de cero (leyendo éstos en el orden natural de izquierda a derecha) de tal línea es 1.
3) Para cada dos líneas consecutivas que no constan exclusivamente de ceros, el primer elemento distinto de cero de la línea superior se encuentra a la izquierda del primer elemento distinto de cero de la línea a la que precede.
4) Cada columna que contiene un primer elemento distinto de cero de alguna línea, tiene en las posiciones restantes 0.

Si la matriz A satisface las condiciones 1), 2) y 3), pero no la 4), se dice que ésta se encuentra en la forma *escalonada*.

EJEMPLO 3 Por ejemplo, las siguientes matrices se encuentran en la forma *escalonada* reducida:

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 6 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

mientras que las matrices

$$\begin{bmatrix} 1 & 0 & 3 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 6 & 2 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 10 \end{bmatrix}$$

se encuentran en la forma *escalonada*.

La propiedad más importante que posee una matriz en la forma *escalonada* reducida es que si ésta representa la matriz aumentada de un sistema de ecuaciones lineales, las soluciones de éste pueden leerse muy fácilmente.

EJEMPLO 4 Por ejemplo, suponga que la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

representa la matriz aumentada de un sistema de ecuaciones lineales, de modo que atendiendo a la correspondencia entre matrices y sistemas de ecuaciones establecida anteriormente (correspondencia entre (2.3) y (2.5)), se ve que el sistema correspondiente es:

$$1 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 4$$
$$0 \cdot x_1 + 1 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = -2$$
$$0 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 + 0 \cdot x_4 = 1$$
$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 1 \cdot x_4 = 4$$

o sea, $x_1 = 4, x_2 = -2, x_3 = 1, x_4 = 4$. Las soluciones del sistema están dadas, pues, en la misma matriz. Por otra parte, si la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

es la matriz aumentada de un sistema de ecuaciones, éste sería entonces

$$x_1 = 3$$
$$x_2 + x_3 = 2$$

Dando cualquier valor a x_3, dígase $x_3 = t$, se obtiene el valor $x_2 = 2 - t$ y entonces el conjunto (infinito) de soluciones del sistema está dado por

$$\{x_1 = 3, x_2 = 2 - t, x_3 = t, t \in \mathbb{R}\}$$

EJEMPLO 5 Como último ejemplo, si se tiene la matriz A en forma *escalonada* reducida

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

se ve que el sistema correspondiente

$$1 \cdot x_1 + 0 \cdot x_2 = 0$$
$$0 \cdot x_1 + 0 \cdot x_2 = 1$$

no tiene solución.

En conclusión, la forma escalonada reducida de una matriz, cuando ésta representa la matriz aumentada de un sistema de ecuaciones, es muy conveniente para de ella leer las soluciones del sistema.

2.2. OPERACIONES ELEMENTALES EN LAS LÍNEAS DE UNA MATRIZ

Al tenerse siempre en mente (por el momento) que las matrices representan sistemas de ecuaciones lineales, se pueden transportar las operaciones elementales en las ecuaciones del sistema, discutidas en la sección anterior, a *operaciones elementales en las líneas de la matriz* (pues cada línea de la matriz corresponde a una ecuación del sistema y viceversa).

Éstas serían pues:

1) Intercambiar dos líneas de la matriz.
2) Multiplicar una línea por una constante distinta de cero.
3) Sustituir una línea por ella misma más k veces otra línea de la matriz.

Se dirá que la matriz B es *equivalente* a la matriz A, lo cual se escribirá $B \sim A$ si se puede obtener B a partir de A, realizando en ésta una secuencia (finita) de operaciones elementales en sus líneas. Véase que, desde el punto de vista de la correspondencia matrices-sistemas de ecuaciones, no se está definiendo más que la equivalencia entre sistemas (es decir, si el sistema que representa la matriz B es equivalente al sistema que representa la matriz A, entonces se dice que B es equivalente a A).

Obsérvese también que esta relación que se ha establecido entre las matrices (B es equivalente a A) es, de hecho, una *relación de equivalencia*, es decir, es **reflexiva, simétrica y transitiva**. En efecto:

1) Toda matriz es equivalente a sí misma (obvio); (reflexibidad).
2) Supóngase que B es equivalente a A. Esto significa que, partiendo de A, se realizan en ésta operaciones elementales que conducen eventualmente a B. Como cada operación elemental tiene su "inversa" que además es del mismo tipo,* se puede, a partir ahora de B realizar las operaciones elementales inversas y llegar así a A. Esto significa que A es equivalente a B (simetría).

Por último:

3) Supóngase que A es equivalente a B y B es equivalente a C. B se obtuvo entonces a partir de C haciendo operaciones elementales. Pero A se obtuvo asimismo partiendo de B. Juntando esta secuencia de operaciones elementales, se puede, a partir de C, obtener A (pasando por B). Esto significa que A es equivalente a C (transitividad).

Se podrá entonces decir, cuando la matriz B es equivalente a la matriz A, que "las matrices A y B son equivalentes" sin temor a confusión.

EJEMPLO 6 Por ejemplo, considere la matriz:

$$A = \begin{bmatrix} 1 & 3 & 2 & 5 \\ 2 & 4 & 1 & 3 \end{bmatrix}$$

Si sustituye la segunda línea de A por ella misma menos 2 veces su primera línea obtenga la matriz

$$A_1 = \begin{bmatrix} 1 & 3 & 2 & 5 \\ 0 & -2 & -3 & -7 \end{bmatrix}$$

La matriz A_1 es entonces equivalente a la matriz A_2.

Multiplique ahora la segunda línea de A_1 por $-\frac{1}{2}$, obteniendo así la matriz

$$A_2 = \begin{bmatrix} 1 & 3 & 2 & 5 \\ 0 & 1 & 3/2 & 7/2 \end{bmatrix}$$

---

*1) Si se han intercambiado las líneas i y j, la operación inversa es en este caso ella misma, esto es, volviendo a intercambiar las líneas i y j se regresa a la matriz original.

2) Si se ha multiplicado una línea por la constante $k \neq 0$, multiplique ahora esa misma línea por la constante k^{-1}.

3) Si la línea r fue sustituida por ella misma más k veces la línea s, sustituya ahora la línea r por ella misma menos k veces la línea s.

La matriz A_2 es equivalente a la matriz A_1 y por tanto es equivalente a A.

Si sustituye ahora la primera línea de A por ella misma menos 3 veces su segunda línea obtenga

$$A_3 = \begin{bmatrix} 1 & 0 & -5/2 & 31/2 \\ 0 & 1 & 3/2 & 7/2 \end{bmatrix}$$

La matriz A_3 es equivalente a la matriz A_2, por tanto equivalente a la matriz A_1 y por tanto equivalente a A.

Observe que A_3 está en forma escalonada reducida. Se dice que la matriz A_3 es "la forma escalonada reducida de la matriz A".

En general se dice que la matriz B es la forma escalonada reducida de la matriz A si

1) B es una matriz en forma escalonada reducida.
2) B es equivalente a A.

2.3. EL MÉTODO DE ELIMINACIÓN GAUSSIANA

Las ideas básicas del método ya fueron discutidas en las subsecciones anteriores. Sólo falta ponerlas en orden: dado el sistema de ecuaciones (2.3) del cual se quieren conocer sus soluciones, escríbase la matriz aumentada asociada al sistema y llévese, por medio de operaciones elementales en sus líneas, a la forma escalonada reducida.* El sistema que representa esta nueva matriz tiene las mismas soluciones que el sistema original (¿por qué?).

El método de *eliminación Gaussiana* indica *cómo*** se puede llevar una matriz dada a su forma escalonada reducida. Las diferentes etapas del método se explicarán con un ejemplo particular.

EJEMPLO 7 Supóngase que se tiene la matriz

$$\begin{bmatrix} 3 & -2 & 2 & 1 & 1 \\ 1 & -1 & -1 & -3 & 0 \\ 2 & 1 & 2 & 4 & 5 \\ 4 & -4 & 1 & 2 & 4 \end{bmatrix}$$

Se hace que el primer elemento distinto de cero de la primera línea sea 1. Esto es posible multiplicando por $\frac{1}{3}$ tal línea. Sin embargo, se puede intercambiar primeramente la primera y segunda líneas (pues esta línea comienza ya con 1) con el objeto de retrasar el manejo de fracciones. Se obtiene pues la matriz

$$\begin{bmatrix} 1 & -1 & -1 & -3 & 0 \\ 3 & -2 & 2 & 1 & 1 \\ 2 & 1 & 2 & 4 & 5 \\ 4 & -4 & 1 & 2 & 4 \end{bmatrix}$$

La siguiente etapa es, por medio de sustituir líneas por ellas mismas más múltiplos de otras, hacer ceros en las posiciones restantes de la columna debajo del 1 logrado en el primer paso. Llamando (L_i) a la línea i, se hace $(L2) \to (L2) - 3(L1)$, $(L3) \to (L3) - 2(L1)$, $(L4) \to (L4) - 4(L1)$. Se obtiene entonces la matriz

$$\begin{bmatrix} 1 & -1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 10 & 1 \\ 0 & 3 & 4 & 10 & 5 \\ 0 & 0 & 5 & 14 & 4 \end{bmatrix}$$

Se concentra ahora la atención en la submatriz señalada en el paso anterior. Se repite de nuevo el proceso. Se tiene ya un 1 como primer elemento, por lo tanto sólo resta hacer ceros debajo de él. Se realiza $(L3) \to (L3) - 3(L2)$ en la matriz anterior obteniéndose así

$$\begin{bmatrix} 1 & -1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 10 & 1 \\ 0 & 0 & -11 & -20 & 2 \\ 0 & 0 & 5 & 14 & 4 \end{bmatrix}$$

Se reitera en la submatriz señalada. Se realiza entonces en la matriz anterior $(L3) \to -\frac{1}{11}(L3)$ para lograr un 1 como primer elemento no nulo de la tercera línea

$$\begin{bmatrix} 1 & -1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 10 & 1 \\ 0 & 0 & 1 & 20/11 & -2/11 \\ 0 & 0 & 5 & 14 & 4 \end{bmatrix}$$

Se hace ahora $(L4) \to (L4) - 5(L3)$

*Se puede demostrar que la forma escalonada reducida de una matriz es única.

**Ciertamente no existe una manera única de proceder para llevar una matriz a su forma escalonada reducida. El método aquí expuesto provee una sistematización de tal procedimiento.

$$\begin{bmatrix} 1 & -1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 10 & 1 \\ 0 & 0 & 1 & 20/11 & -2/11 \\ 0 & 0 & 0 & 54/11 & 54/11 \end{bmatrix}$$

Para tener un 1 como primer elemento no nulo de la cuarta línea se realiza $(L4) \rightarrow \frac{11}{54}(L4)$

$$\begin{bmatrix} 1 & -1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 10 & 1 \\ 0 & 0 & 1 & 20/11 & -2/11 \\ 0 & 0 & 0 & 1 & 11 \end{bmatrix}$$

Obsérvese que esta matriz se encuentra ya en la forma *escalonada*. Para llegar a la forma *escalonada* reducida, se comienza con el *último uno logrado* (en el paso anterior) y se procede a volver ceros las posiciones restantes (encima de él). Para esto se realiza $(L1) \rightarrow -(L1) + 3(L4)$, $(L2) \rightarrow (L2) - 10(L4)$ y $(L3) \rightarrow (L3) - \frac{20}{11}(L4)$

$$\begin{bmatrix} 1 & -1 & -1 & 0 & 3 \\ 0 & 1 & 5 & 0 & -9 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Ahora con el siguiente uno se hace lo mismo: se realiza $(L1) \rightarrow (L1) + (L3)$, $(L2) \rightarrow (L2) - 5(L3)$

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Finalmente se efectúa $(L1) \rightarrow (L1) + (L2)$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

y el sistema que representa esta matriz es entonces $x_1 = 2$, $x_2 = 1$, $x_3 = -2$, $x_4 = 1$ que es la solución del sistema. (Compárese el procedimiento con el que se empleó para resolver este mismo sistema en la página 27).

Véanse un par de ejemplos más.

EJEMPLO 8

Se quiere resolver el sistema

$$\begin{aligned} 2x_1 + 6x_2 + 14x_3 + 4x_4 &= 4 \\ x_1 - 2x_2 + x_3 - 4x_4 &= 1 \\ x_1 - 12x_2 - 11x_3 - 16x_4 &= 15 \end{aligned}$$

La matriz aumentada es:

$$\begin{bmatrix} 2 & 6 & 14 & 4 & 4 \\ 1 & -2 & 1 & -4 & 1 \\ 1 & -12 & -11 & -16 & 15 \end{bmatrix}$$

Se exponen en seguida los pasos para llegar a la forma escalonada reducida de esta matriz. En cada paso se indican las operaciones elementales que se efectúan

$$\begin{bmatrix} 2 & 6 & 14 & 4 & 4 \\ 1 & -2 & 1 & -4 & 1 \\ 1 & -12 & -11 & -16 & 15 \end{bmatrix} (L1) \rightarrow \frac{1}{2}(L1) \begin{bmatrix} 1 & 3 & 7 & 2 & 2 \\ 1 & -2 & 1 & -4 & 1 \\ 1 & -12 & -11 & -16 & 15 \end{bmatrix} \begin{matrix} (L2) \rightarrow (L2) - (L1) \\ (L3) \rightarrow (L3) - (L1) \end{matrix}$$

$$\begin{bmatrix} 1 & 3 & 7 & 2 & 2 \\ 0 & -5 & -6 & -6 & -1 \\ 0 & -15 & -18 & -18 & 13 \end{bmatrix} (L2) \rightarrow -\frac{1}{5}(L2) \begin{bmatrix} 1 & 3 & 7 & 2 & 2 \\ 0 & 1 & 6/5 & 6/5 & 1/5 \\ 0 & -15 & -18 & -18 & 13 \end{bmatrix} (L3) \rightarrow (L3) + 15(L2)$$

$$\begin{bmatrix} 1 & 3 & 7 & 2 & 2 \\ 0 & 1 & 6/5 & 6/5 & 1/5 \\ 0 & 0 & 0 & 0 & 16 \end{bmatrix}$$

A pesar de que no se ha llegado aún a la forma escalonada reducida de la matriz, se puede observar que en este último paso está contenida toda la información que se necesita: la ecuación del sistema que representa esta última matriz (equivalente a la primera) es:

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 16$$

Es claro, entonces, que el sistema es inconsistente.

EJEMPLO 9

Un último ejemplo. Se quiere resolver el sistema

$$
\begin{aligned}
2x_1 + 2x_2 + 4x_3 + \quad + \quad &= 10 \\
x_2 + 2x_3 + 2x_4 + 12x_5 &= 0 \\
x_1 + x_2 + 2x_3 + 5x_4 + 30x_5 &= 0 \\
x_1 + x_2 + 2x_3 + 2x_4 + 12x_5 &= 3
\end{aligned}
$$

El procedimiento de eliminación se ve como sigue:

$$
\begin{bmatrix} 2 & 2 & 4 & 0 & 0 & 10 \\ 0 & 1 & 2 & 2 & 12 & 0 \\ 1 & 1 & 2 & 5 & 30 & 0 \\ 1 & 1 & 2 & 2 & 12 & 3 \end{bmatrix}
(L1)\tilde{\to}\frac{1}{2}(L1)
\begin{bmatrix} 1 & 1 & 2 & 0 & 0 & 5 \\ 0 & 1 & 2 & 2 & 12 & 0 \\ 1 & 1 & 2 & 5 & 30 & 0 \\ 1 & 1 & 2 & 2 & 12 & 3 \end{bmatrix}
$$

$$
\begin{matrix} (L3)\tilde{\to}(L3)-(L1) \\ (L4)\to(L4)-(L1) \end{matrix}
\begin{bmatrix} 1 & 1 & 2 & 0 & 0 & 5 \\ 0 & 1 & 2 & 2 & 12 & 0 \\ 0 & 0 & 0 & 5 & 30 & -5 \\ 0 & 0 & 0 & 2 & 12 & -2 \end{bmatrix}
(L3)\tilde{\to}\frac{1}{5}(L3)
\begin{bmatrix} 1 & 1 & 2 & 0 & 0 & 5 \\ 0 & 1 & 2 & 2 & 12 & 0 \\ 0 & 0 & 0 & 1 & 6 & -1 \\ 0 & 0 & 0 & 2 & 12 & -2 \end{bmatrix}
$$

$$
(L4)\tilde{\to}(L4)-2(L3)
\begin{bmatrix} 1 & 1 & 2 & 0 & 0 & 5 \\ 0 & 1 & 2 & 2 & 12 & 0 \\ 0 & 0 & 0 & 1 & 6 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
(L2)\tilde{\to}(L2)-2(L3)
\begin{bmatrix} 1 & 1 & 2 & 0 & 0 & 5 \\ 0 & 1 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 6 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

$$
(L1)\tilde{\to}(L1)-(L2)
\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 3 \\ 0 & 1 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 6 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

El sistema equivalente al original es:

$$
\begin{aligned}
x_1 &= 3 \\
x_2 + 2x_3 &= 2 \\
x_4 + 6x_5 &= -1
\end{aligned}
$$

Véase entonces que se pueden dar valores arbitrarios a x_3 y a x_5, dígase $x_3 = r$, $x_5 = s$, y quedan así determinados los valores de x_2 y x_4, $x_2 = 2 - 2r$, $x_4 = -1-6s$.

El conjunto solución del sistema es entonces

$$\{x_1 = 3, x_2 = 2-2r, x_3 = r, x_4 = -1-6s, x_5 = s, \quad (r, s \in \mathbb{R})\}$$

Antes de pasar a la siguiente subsección, se quisiera en este momento sacar un poco más de provecho al método de eliminación Gaussiana discutido aquí para resolver otros tipos de problemas relacionados con sistemas de ecuaciones que aparecerán posteriormente en el libro.

Considérese el sistema de ecuaciones

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\
&\ldots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m
\end{aligned}
$$

Supóngase dados los coeficientes a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, teniendo indeterminados los términos independientes $b_1, b_2, \ldots, b_m$. Se trata de ver el comportamiento de las soluciones del sistema de acuerdo a los posibles valores de los términos independientes. El método de eliminación Gaussiana ayuda a dar respuesta a este problema.

EJEMPLO 10 Si se considera el sistema

$$
\begin{aligned}
3x_1 + x_2 + 5x_3 &= b_1 \\
2x_1 + 4x_2 + x_3 &= b_2 \\
-4x_1 + 2x_2 - 9x_3 &= b_3
\end{aligned}
$$

y uno se pregunta, por ejemplo, para qué valores de b_1, b_2 y b_3 el sistema tiene solución, se puede proceder con el método de eliminación de la siguiente manera: llevando la matriz aumentada a su forma *escalonada* reducida

$$
\begin{bmatrix} 3 & 1 & 5 & b_1 \\ 2 & 4 & 1 & b_2 \\ -4 & 2 & -9 & b_3 \end{bmatrix}
(L1)\tilde{\to}\frac{1}{3}(L1)
\begin{bmatrix} 1 & 1/3 & 5/3 & b_1/3 \\ 2 & 4 & 1 & b_2 \\ -4 & 2 & -9 & b_3 \end{bmatrix}
\begin{matrix} (L2)\tilde{\to}(L2)-2(L1) \\ (L3)\to(L3)+4(L1) \end{matrix}
$$

$$
\begin{bmatrix} 1 & 1/3 & 5/3 & b_1/3 \\ 0 & 10/3 & -7/3 & b_2-2b_1/3 \\ 0 & 10/3 & -7/3 & b_3+4b_1/3 \end{bmatrix}
(L2)\tilde{\to}\frac{3}{10}(L2)
\begin{bmatrix} 1 & 1/3 & 5/3 & b_1/3 \\ 0 & 1 & -7/10 & \dfrac{3b_2-2b_1}{10} \\ 0 & 10/3 & -7/3 & b_3+\dfrac{4b_1}{3} \end{bmatrix}
$$

$$
(L3)\tilde{\to}(L3)-\frac{10}{3}(L2)
\begin{bmatrix} 1 & 1/3 & 5/3 & b_1/3 \\ 0 & 1 & -7/10 & \dfrac{3b_2-2b_1}{10} \\ 0 & 0 & 0 & b_3-b_2+2b_1 \end{bmatrix}
(L1)\tilde{\to}(L1)-\frac{1}{3}(L2)
$$

$$
\begin{bmatrix} 1 & 0 & 57/30 & \dfrac{4b_1-b_2}{10} \\ 0 & 1 & -7/10 & \dfrac{3b_2-2b_1}{10} \\ 0 & 0 & 0 & b_3-b_2+2b_1 \end{bmatrix}
$$

véase entonces que, a menos que $b_3 - b_2 + 2b_1 = 0$, el sistema no tiene solución.

En el caso $b_3 - b_2 + 2b_1 = 0$, las soluciones del sistema serían

$$x_1 = \frac{4b_1 - b_2}{10} - \frac{57}{30}r$$

$$x_2 = \frac{3b_2 - 2b_1}{10} + \frac{7}{10}r \qquad r \in \mathbf{R}$$

$$x_3 = r$$

Otro tipo de problema interesante que puede resolverse usando el método de eliminación es el siguiente: supóngase dado un sistema de ecuaciones en el que existe un parámetro indeterminado dentro de los coeficientes del sistema. Se pregunta también sobre el comportamiento de las soluciones en función de los valores de este parámetro.

EJEMPLO 11 Por ejemplo, considere el sistema

$$\lambda x_1 + x_2 + x_3 + x_4 = 1$$

$$x_1 + \lambda x_2 + x_3 + x_4 = 2$$

$$x_1 + x_2 + \lambda x_3 + x_4 = 3$$

$$x_1 + x_2 + x_3 + \lambda x_4 = 4$$

Realice el proceso de llevar la matriz aumentada del sistema a su forma *escalonada* reducida

$$\begin{bmatrix} \lambda & 1 & 1 & 1 & 1 \\ 1 & \lambda & 1 & 1 & 2 \\ 1 & 1 & \lambda & 1 & 3 \\ 1 & 1 & 1 & \lambda & 4 \end{bmatrix} (L1) \leftrightarrow (L4) \begin{bmatrix} 1 & 1 & 1 & \lambda & 4 \\ 1 & \lambda & 1 & 1 & 2 \\ 1 & 1 & \lambda & 1 & 3 \\ \lambda & 1 & 1 & 1 & 1 \end{bmatrix} \begin{matrix} (L2) \to (L2) - (L1) \\ (L3) \to (L3) - (L1) \\ (L4) \to (L4) - \lambda (L1) \end{matrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & \lambda & 4 \\ 0 & \lambda - 1 & 0 & 1 - \lambda & -2 \\ 0 & 0 & \lambda - 1 & 1 - \lambda & -1 \\ 0 & 1 - \lambda & 1 - \lambda & 1 - \lambda^2 & 1 - 4\lambda \end{bmatrix}$$

El siguiente paso en nuestro proceso de eliminación sería dividir la segunda línea por $\lambda - 1$. Pero esto se puede hacer sólo cuando $\lambda \neq 1$. Si $\lambda = 1$ se obtendría la matriz

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -3 \end{bmatrix}$$

que representa a un sistema sin solución. Por lo tanto, se concluye que si $\lambda = 1$ el sistema no tiene solución.

Supóngase ahora que $\lambda \neq 1$ y siga adelante en la eliminación Se realiza $(L_i) \to \frac{1}{\lambda - 1}(L_i)$, $i = 2, 3, 4$, para obtener

$$\begin{bmatrix} 1 & 1 & 1 & \lambda & 4 \\ 0 & 1 & 0 & -1 & -2/\lambda - 1 \\ 0 & 0 & 1 & -1 & -1/\lambda - 1 \\ 0 & -1 & -1 & -(1+\lambda) & \frac{1-4\lambda}{\lambda - 1} \end{bmatrix} (L4) \to (L4) + (L2)$$

$$\begin{bmatrix} 1 & 1 & 1 & \lambda & 4 \\ 0 & 1 & 0 & -1 & -\frac{2}{\lambda - 1} \\ 0 & 0 & 1 & -1 & -\frac{1}{\lambda - 1} \\ 0 & 0 & -1 & -(2+\lambda) & -\frac{1+4\lambda}{\lambda - 1} \end{bmatrix} (L4) \to (L4) + (L3)$$

$$\begin{bmatrix} 1 & 1 & 1 & \lambda & 4 \\ 0 & 1 & 0 & -1 & -\frac{2}{\lambda - 1} \\ 0 & 0 & 1 & -1 & -\frac{1}{\lambda - 1} \\ 0 & 0 & 0 & -(3+\lambda) & -\frac{2+4\lambda}{\lambda - 1} \end{bmatrix}$$

Nuevamente, el siguiente paso sería dividir por $-(3 + \lambda)$ la cuarta línea, lo cual se puede hacer cuando $\lambda \neq -3$. Si $\lambda = -3$, la última ecuación que representa la matriz obtenida sería $0 = -\frac{5}{2}$ por lo que el sistema no tendría solución. Se concluye entonces: si $\lambda = -3$ el sistema tampoco tiene solución. Sea pues $\lambda \neq -3$ y continúe

$$\begin{bmatrix} 1 & 1 & 1 & \lambda & 4 \\ 0 & 1 & 0 & -1 & -\frac{2}{\lambda - 1} \\ 0 & 0 & 1 & -1 & -\frac{1}{\lambda - 1} \\ 0 & 0 & 0 & 1 & -\frac{2+4\lambda}{(\lambda - 1)(3+\lambda)} \end{bmatrix} \begin{matrix} (L1) \to (L1) - \lambda (L4) \\ (L2) \to (L2) + (L4) \\ (L3) \to (L3) + (L4) \end{matrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & \frac{6\lambda-12}{(\lambda-1)(\lambda+3)} \\ 0 & 1 & 0 & 0 & \frac{2\lambda-4}{(\lambda-1)(\lambda+3)} \\ 0 & 0 & 1 & 0 & \frac{3\lambda-1}{(\lambda-1)(\lambda+3)} \\ 0 & 0 & 0 & 1 & \frac{2+4\lambda}{(\lambda-1)(\lambda+3)} \end{bmatrix} \sim \quad (L1) \to (L1) - (L3)$$

$$\begin{bmatrix} 1 & 1 & 0 & 0 & \frac{3\lambda-11}{(\lambda-1)(\lambda+3)} \\ 0 & 1 & 0 & 0 & \frac{2\lambda-4}{(\lambda-1)(\lambda+3)} \\ 0 & 0 & 1 & 0 & \frac{3\lambda-1}{(\lambda-1)(\lambda+3)} \\ 0 & 0 & 0 & 1 & \frac{2+4\lambda}{(\lambda-1)(\lambda+3)} \end{bmatrix} \sim \quad (L1) \to (L1) - (L2)$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \frac{\lambda-7}{(\lambda-1)(\lambda+3)} \\ 0 & 1 & 0 & 0 & \frac{2\lambda-4}{(\lambda-1)(\lambda+3)} \\ 0 & 0 & 1 & 0 & \frac{3\lambda-1}{(\lambda-1)(\lambda+3)} \\ 0 & 0 & 0 & 1 & \frac{2+4\lambda}{(\lambda-1)(\lambda+3)} \end{bmatrix}$$

En conclusión, se tienen dos casos:

1) Si $\lambda = 1, -3$, el sistema no tiene solución.
2) Si $\lambda \neq 1, -3$, el sistema tiene una única solución, a saber

$$x_1 = \frac{\lambda-7}{(\lambda-1)(\lambda+3)}, \; x_2 = \frac{2\lambda-4}{(\lambda-1)(\lambda+3)}, \; x_3 = \frac{3\lambda-1}{(\lambda-1)(\lambda+3)}, \; x_4 = \frac{2+4\lambda}{(\lambda-1)(\lambda+3)}$$

NOTA: A lo largo del material que se lleva analizado hasta este momento, se ha podido observar en los ejemplos que el conjunto solución de un sistema de ecuaciones lineales puede ser:

1) **Vacío.** En este caso se dice que el sistema no tiene solución.
2) **Finito.** Formando por *un* valor para cada incógnita. En este caso se dice que el sistema tiene solución única.
3) **Infinito.** Formado por una infinidad de valores para al menos una de las incógnitas. En este caso se dice que el sistema tiene una infinidad de soluciones.

Estos casos corresponden geométricamente, en el caso de un sistema de 2 ecuaciones con 2 incógnitas (tal caso como se observó en la introducción al capítulo) a: 1) las rectas que representan las ecuaciones son paralelas; 2) las rectas se cortan en un punto, 3) las rectas coinciden, respectivamente.

Se usará también la terminología: *sistema inconsistente* para un sistema cuyo conjunto solución sea vacío y *sistema consistente* para uno cuyo conjunto solución sea **no** vacío.

Se puede demostrar (se hará) que, dado un sistema de ecuaciones lineales, su conjunto solución *siempre* será de uno de estos 3 tipos.

Dado un sistema de ecuaciones en general *no se puede saber a priori* de qué tipo es su conjunto solución. Sin embargo, hay un caso especial (importante) en el que sí se puede saber en base al número de ecuaciones y de incógnitas del sistema. Éste es el caso de un sistema homogéneo.

2.4. SISTEMAS HOMOGÉNEOS DE ECUACIONES LINEALES

Considérese el sistema homogéneo

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0 \\ &\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0 \end{aligned} \qquad (2.6)$$

Obsérvese que este sistema *siempre será consistente*, pues al menos posee la solución $x_1 = x_2 = \ldots = x_n = 0$, llamada *solución trivial.** Aún más, se puede saber (sin resolverse) bajo qué condiciones posee, además de la solución trivial, otras soluciones (y entonces, según lo señalado en la nota final de la subsección (2.3), en tal caso el sistema tendrá una infinidad de soluciones). Antes de enunciar y probar el teorema que enuncia con precisión este hecho, véanse un par de ejemplos.

EJEMPLO 12 Considere primeramente el sistema homogéneo de 4 ecuaciones con 3 incógnitas

$$\begin{aligned} 5x_1 + 7x_2 - 9x_3 &= 0 \\ x_1 + x_2 - x_3 &= 0 \\ 3x_1 + x_2 + x_3 &= 0 \\ -2x_1 - 4x_2 + 6x_3 &= 0 \end{aligned}$$

*En el caso $m = n = 2$ se puede ver geométricamente este hecho. Una ecuación lineal homogénea en dos variables representa geométricamente una recta que pasa por el origen. Por lo tanto, dado un sistema de dos ecuaciones lineales homogéneas con dos incógnitas, las rectas que estas ecuaciones representan se cortan *al menos* en el origen (solución trivial) quedando también la posibilidad de que tales rectas coincidan (infinidad de soluciones).

Para obtener las soluciones del sistema, use el método de eliminación Gaussiana, observando que en este caso no hay necesidad de considerar la matriz aumentada del sistema, bastando sólo con la matriz del sistema. Esto es porque las operaciones elementales que realice en tal matriz no alterarán en absoluto la columna de los términos independientes, que consta sólo de ceros.

Hecho este convenio, proceda a la eliminación (los pasos intermedios quedan a su cargo)

$$\begin{bmatrix} 5 & 7 & -9 \\ 1 & 1 & -1 \\ 3 & 1 & 1 \\ -2 & -4 & 6 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

de modo que el sistema dado es equivalente a

$$x_1 + x_3 = 0$$

$$x_2 - 2x_3 = 0$$

y entonces, el conjunto solución será

$$\{x_1 = -t,\ x_2 = 2t,\ x_3 = t,\ t \in \mathbf{R}\}$$

EJEMPLO 13 Resuelva ahora el siguiente sistema homogéneo de 3 ecuaciones con 4 incógnitas:

$$x_1 + 2x_2 - 3x_3 + x_4 = 0$$

$$2x_1 + 2x_3 - x_4 = 0$$

$$x_1 - x_2 + x_3 - 2x_4 = 0$$

Haciendo las operaciones correspondientes obtenga

$$\begin{bmatrix} 1 & 2 & -3 & 1 \\ 2 & 0 & 2 & -1 \\ 1 & -1 & 1 & -2 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & -7/8 \\ 0 & 1 & 0 & 3/2 \\ 0 & 0 & 1 & 3/8 \end{bmatrix}$$

de modo que el sistema

$$x_1 - \frac{7}{8}x_4 = 0$$

$$x_2 + \frac{3}{2}x_4 = 0$$

$$x_3 + \frac{3}{8}x_4 = 0$$

es equivalente al original y entonces el conjunto solución es:

$$\left\{x_1 = \frac{7}{8}t,\ x_2 = -\frac{3}{2}t,\ x_3 = -\frac{3}{8}t,\ x_4 = t,\ t \in \mathbf{R}\right\}$$

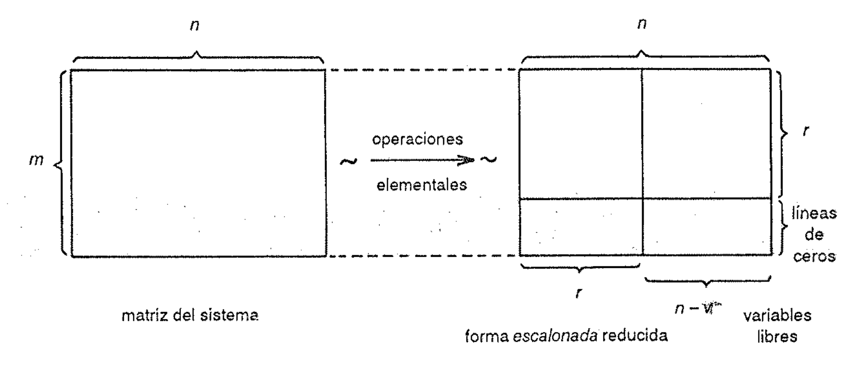
Analice con detenimiento este último ejemplo. La matriz del sistema es una matriz de 3 líneas (una por cada ecuación) y 4 columnas (correspondientes a las 4 incógnitas del sistema). Al llevar esta matriz a su forma escalonada reducida véase que, por ser menor el número de líneas que el de columnas, se debe esperar que *al menos* una de las columnas de la matriz no tenga su primer elemento distinto de cero (un 1) de línea alguna. El peor de los casos (el que ocurrió en el ejemplo) es cuando se tiene 4 - 3 = 1 columna no formada por ceros y unos. Esta columna corresponde a la cuarta incógnita, que quedó entonces como *variable libre*. Como es de esperarse, una situación similar ocurrirá siempre que se tenga un sistema homogéneo con más incógnitas que ecuaciones. El siguiente teorema establece en general este resultado:

TEOREMA 2.1 Un sistema homogéneo de ecuaciones lineales con más incógnitas que ecuaciones tendrá siempre una infinidad de soluciones.

DEMOSTRACIÓN Supóngase que el sistema tiene m ecuaciones y n incógnitas. La hipótesis del teorema es que $m < n$. La matriz del sistema es una matriz *rectangular horizontal*. Supóngase que, después de llevar esta matriz a su forma escalonada reducida, se obtuvieron r líneas no nulas. Es claro que $r \leq m$, y por lo tanto $r < n$. Entonces, existen $k = n - r$ variables que quedan libres en el sistema, esto es, que éste posee una infinidad de soluciones.

Q.E.D.

La siguiente representación esquemática ayuda a comprender el teorema anterior.



Obsérvese, sin embargo, que el teorema no dice nada acerca de las soluciones para un sistema con menor o igual número de incógnitas que ecuaciones, pudiéndose presentar siempre los dos casos (sólo la solución trivial o infinidad de soluciones).

Ya en el primer ejemplo visto en esta subsección se tenía un sistema con menos incógnitas que ecuaciones y sin embargo el sistema también posee una infinidad de soluciones.

EJERCICIOS (SECCIÓN 2, CAPÍTULO 1)

1. Para cada uno de los sistemas de ecuaciones dados a continuación, escriba la matriz del sistema, así como la matriz aumentada del mismo.

a) $2x_1 + x_2 = 0$
$3x_1 - x_2 = 5$

b) $2x_1 + 3x_2 - 4x_3 = 7$
$3x_1 - 4x_2 + 7x_3 = 4$
$6x - 7x_2 + 8x_3 = 4$
$8x_1 - 4x_2 + 7x_3 = 9$

c) $x_1 = 0$
$x_2 = 0$
$x_3 = 1$
$x_4 = 0$

d) $x_1 + x_2 = 0$
$x_2 + x_3 = 0$
$x_3 + x_4 = 0$
$x_4 + x_5 = 0$
$x_5 + x_6 = 0$

2. Cada una de las siguientes matrices es la matriz aumentada de un sistema de ecuaciones lineales. Escriba el sistema correspondiente a cada matriz.

a) $\begin{bmatrix} 1 & 7 & 8 & 0 \\ 2 & 1 & 4 & -1 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 1 & 3 \\ 1 & 0 & 1 & 0 & 4 \\ 0 & 1 & 0 & 1 & 5 \end{bmatrix}$

d) $\begin{bmatrix} 0 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 2 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 2 \\ 1 & 0 & 0 & 0 & 3 \end{bmatrix}$

e) $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

3. De las siguientes matrices, diga cuáles están en la forma escalonada, cuáles en la forma escalonada reducida o cuáles no están en ninguna de estas dos formas.

a) $\begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

d) $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 1 & 2 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

e) $\begin{bmatrix} 0 & 0 & 1 & 2 & 3 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

4. Suponga que las matrices del ejercicio anterior representan matrices aumentadas de sistemas de ecuaciones lineales. Determine las soluciones de estos sistemas.

5. En cada uno de los siguientes incisos, la matriz A_2 se ha obtenido de la matriz A_1 por medio de una operación elemental en las líneas de esta última. Identifique tal operación elemental y describa qué operación habrá de realizar en las líneas de A_2 para recuperar así A_1 (la operación elemental inversa correspondiente).

a) $A_1 = \begin{bmatrix} 1 & 3 & 2 & 1 \\ 4 & 5 & 1 & 2 \\ 0 & 3 & 1 & 4 \end{bmatrix}$ $\quad A_2 = \begin{bmatrix} 4 & 5 & 1 & 2 \\ 1 & 3 & 2 & 1 \\ 0 & 3 & 1 & 4 \end{bmatrix}$

b) $A_1 = \begin{bmatrix} 1 & 3 & 1 & 2 & 1 \\ 2 & 1 & 4 & 1 & 3 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 1 & 3 & 1 & 0 \end{bmatrix}$ $\quad A_2 = \begin{bmatrix} 1 & 3 & 1 & 2 & 1 \\ 2 & 1 & 4 & 1 & 3 \\ 3 & 9 & 4 & 8 & 6 \\ 0 & 1 & 3 & 1 & 0 \end{bmatrix}$

c) $A_1 = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 3 & 1 & 2 & 1 \\ 4 & 1 & 0 & 3 \end{bmatrix}$ $\quad A_2 = \begin{bmatrix} 7 & 14 & 7 & 21 \\ 3 & 1 & 2 & 1 \\ 4 & 1 & 0 & 3 \end{bmatrix}$

6. Describa todas las posibles formas escalonadas reducidas para una matriz cuadrada de orden 2.

7. ¿Bajo qué condiciones la forma escalonada reducida de la matriz

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

es

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}?$$

8. Describa todas las posibles formas escalonadas reducidas para una matriz cuadrada de orden 3.

① 9. En esta sección se presentaron 3 tipos de operaciones elementales en las líneas de una matriz:

1) Intercambiar dos líneas de la matriz.
2) Multiplicar una línea de la matriz por una constante distinta de cero.
3) Sustituir una línea por ella misma más k veces otra línea de la matriz.

Demuestre que la operación del tipo 1 puede ser sustituida por una secuencia adecuada de operaciones del tipo 2 y 3.

En la matriz

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

intercambie la primera y tercera líneas, usando solamente operaciones elementales del tipo 2 y 3.

10. Sean A y B dos matrices del mismo orden. Suponga que estas matrices son las matrices aumentadas de los sistemas de ecuaciones s_A y s_B respectivamente.

Demuestre que la matriz A es equivalente a la matriz B si, y sólo si el sistema s_A y el sistema s_B son equivalentes.

11. Use el método de eliminación de Gauss-Jordan para resolver los siguientes sistemas de ecuaciones lineales.

a) $$\begin{aligned} 2x_1 + 3x_2 &= 5 \\ x_1 - 7x_2 &= 4 \end{aligned}$$

b) $$\begin{aligned} 3x_1 + x_2 &= 1 \\ x_1 + x_2 &= -1 \end{aligned}$$

c) $$\begin{aligned} 5x_1 + 7x_2 - x_3 &= 0 \\ 2x_1 + 10x_2 - 2x_3 &= 4 \\ x_1 + x_2 + x_3 &= 0 \end{aligned}$$

d) $$\begin{aligned} x_1 + 2x_2 + x_3 &= 4 \\ 5x_1 + x_2 + 6x_3 &= 0 \\ 2x_1 + 3x_2 + 3x_3 &= 4 \end{aligned}$$

e) $$\begin{aligned} 5x_1 - 7x_2 + 8x_3 - 2x_4 &= 10 \\ x_1 + 3x_2 - 7x_3 + 3x_4 &= -10 \\ 2x_1 - 5x_2 + 3x_3 - 3x_4 &= 6 \\ 3x_1 + 6x_2 + x_3 + 7x_4 &= -6 \end{aligned}$$

f) $$\begin{aligned} 2x_1 + 3x_2 - x_3 + 2x_4 &= 3 \\ x_1 + 2x_2 + x_3 - 3x_4 &= 1 \\ 2x_1 + x_2 - 6x_3 + x_4 &= -1 \end{aligned}$$

g) $$\begin{aligned} 3x_1 + 2x_2 - x_3 + x_4 &= 7 \\ x_1 - x_2 + x_3 - 2x_4 &= 5 \\ 4x_1 + x_2 - x_4 &= 6 \end{aligned}$$

h) $$\begin{aligned} 2x_1 + 3x_2 - x_3 + 6x_4 - x_5 &= 1 \\ x_1 + 2x_2 + x_3 - 2x_4 + 3x_5 &= -3 \\ 2x_1 + x_2 - 6x_3 + 5x_4 + 3x_5 &= 2 \end{aligned}$$

i) $$\begin{aligned} 3x_1 + 2x_2 - x_3 - x_4 + x_5 &= 0 \\ x_1 - x_2 + 6x_3 - x_4 - 3x_5 &= 0 \\ 2x_1 + x_2 - 7x_3 + x_4 + x_5 &= 0 \end{aligned}$$

j) $$\begin{aligned} 2x_1 + 3x_2 - x_3 + 5x_4 - 2x_5 &= 4 \\ x_1 + 2x_2 + x_3 + 3x_4 - x_5 &= -1 \\ 2x_1 + x_2 - 6x_3 + 7x_4 + 3x_5 &= 1 \\ 5x_1 + 11x_2 + 7x_3 + 12x_4 - 10x_5 &= 4 \end{aligned}$$

k) $$\begin{aligned} x_1 + 2x_2 + 3x_3 + 3x_4 + 5x_5 &= 15 \\ x_2 + 2x_3 + 3x_4 + 4x_5 &= 10 \\ x_3 + 2x_4 + 3x_5 &= 6 \\ x_4 + 2x_5 &= 3 \\ x_5 &= 1 \end{aligned}$$

l) $$\begin{aligned} 2x_1 - 4x_2 + 8x_3 - 14x_4 &= 0 \\ 3x_1 - x_2 + 2x_3 - 7x_4 &= 0 \\ 2x_1 + 3x_2 - x_3 + 5x_4 &= 0 \\ 4x_1 + x_2 - 3x_3 + 6x_4 &= 0 \end{aligned}$$

m) $$\begin{aligned} 3x_1 + 2x_2 - 2x_3 + x_4 - x_5 &= 1 \\ 5x_1 + 5x_3 - 2x_4 + 4x_5 &= 1 \\ x_1 + 3x_2 - 2x_3 + 5x_4 - 7x_5 &= 2 \\ 4x_1 + x_2 + 5x_3 + 3x_5 &= 3 \end{aligned}$$

n) $$\begin{aligned} x_1 - 2x_2 + x_3 + 3x_4 - 3x_5 &= 0 \\ x_1 + 7x_2 - 5x_3 - 5x_4 + 5x_5 &= 0 \\ 2x_1 + x_2 - x_3 - 4x_4 + 4x_5 &= 0 \\ 3x_1 - x_2 - 2x_3 + x_4 - x_5 &= 0 \end{aligned}$$

o) $$\begin{aligned} 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 &= -2 \\ 4x_1 + 3x_2 + 2x_3 + 2x_4 - 2x_5 &= 5 \\ 5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 &= 12 \\ 5x_1 + 5x_2 + 5x_3 + 5x_4 + 5x_5 &= 35 \end{aligned}$$

p) $$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 + x_6 &= 3 \\ 2x_1 + x_2 + 3x_3 + x_4 - 5x_5 + x_6 &= 0 \\ x_1 + 2x_2 + x_3 - 2x_4 + 2x_5 - x_6 &= 4 \\ x_1 - x_2 + x_3 + 4x_4 + 4x_5 + 3x_6 &= 6 \\ x_1 - x_2 - x_3 + x_4 + x_5 - x_6 &= 1 \\ x_1 + x_2 - x_3 - x_4 - x_5 - x_6 &= -1 \end{aligned}$$

① 12. Resuelva el siguiente sistema de n ecuaciones con n incógnitas:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \end{aligned}$$

en donde $a_{ij} = i + j$, $b_i = i$, $i, j = 1, 2, \ldots, n$.

13. Sean $p(x)$ y $q(x)$ dos polinomios del mismo grado, diga

$$\begin{aligned} p(x) &= a_0 + a_1x + \ldots + a_nx^n \\ q(x) &= b_0 + b_1x + \ldots + b_nx^n \end{aligned}$$

Se dice que el polinomio $p(x)$ es igual al polinomio $q(x)$, lo cual se escribe $p(x) = q(x)$, si $a_i = b_i$ $i = 0, 1, \ldots, n$.

a) Determine los valores de a y b para que se cumpla la igualdad de polinomios

$$x^2 + (2a + b)x + 3a - b = x^2 + 5x - 1$$

b) Determine los valores de a y b para que se cumpla la igualdad

$$\frac{2x+1}{(x-1)(x-2)} = \frac{a}{x-1} + \frac{b}{x-2}$$

(*Sugerencia*: multiplique ambos miembros por $(x - 1)(x - 2)$).

c) Determine los valores de a, b y c para que se cumpla la igualdad

$$\frac{3x^2 + 2x + 1}{x^3 - 1} = \frac{a}{x-1} + \frac{bx + c}{x^2 + x + 1}$$

(*Sugerencia*: multiplique ambos miembros por $x^3 - 1$).

d) Determine los valores de a, b, c y d para que se cumpla la igualdad

$$\frac{x^3 + x^2 + x + 1}{x^4 - 1} = \frac{a}{x-1} + \frac{b}{x+1} + \frac{cx + d}{x^2 + 1}$$

(*Sugerencia*: multiplique ambos miembros por $x^4 - 1$).

14. ¿Para qué valores de a, el sistema

$$\begin{aligned} ax_1 + x_2 &= 1 \\ x_1 + ax_2 &= 2 \end{aligned}$$

tiene una solución única, ¿cuál es esta solución?

15. Considere el sistema de ecuaciones lineales

$$\begin{aligned} ax_1 - 3ax_2 &= 1 \\ 2ax_1 + (6a + 1)x_2 &= 2 \end{aligned}$$

Determine el valor de a para el cual

a) el sistema se tiene una solución única

b) el sistema no tiene soluciones

16. Considere el sistema de ecuaciones lineales

$$\begin{aligned} x_1 + x_2 + (a^2 - 1)x &= 1 \\ x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_1 + 5x_2 + x_3 &= b \end{aligned}$$

Clasifique las diferentes posibilidades para el conjunto solución de este sistema en términos de los parámetros a y b.

17. Considere el sistema de ecuaciones lineales

$$\begin{aligned} -2ax_1 - 5x_2 + 2x_3 &= 1 \\ 2x_1 + x_2 + ax_3 &= 1 \\ 2x_1 + 3x_2 + ax_3 &= 2 \end{aligned}$$

Demuestre que independientemente del valor de a, el sistema posee una solución única. Determine tal solución.

18. Clasifique las diferentes posibilidades para el conjunto solución del sistema siguiente, en términos de los parámetros δ, a, b, y c

$$\begin{aligned} x_1 + x_2 + x_3 &= a \\ \delta x_1 + 2x_2 + x_3 &= b \\ \delta^2 x_1 + 4x_2 + x_3 &= c \end{aligned}$$

19. Para los siguientes sistemas de ecuaciones, determine los valores a y b para que: 1) el sistema tenga una solución única; 2) el sistema no tenga solución; 3) el sistema tenga una infinidad de soluciones

1) $\begin{aligned} 2x_1 + 3x_2 + x_3 &= 5 \\ 3x_1 - x_2 + ax_3 &= 2 \\ x_1 + 7x_2 - 6x_3 &= b \end{aligned}$

2) $\begin{aligned} x_1 + x_2 - 4x_3 &= 0 \\ 2x_1 + 3x_2 + x_3 &= 1 \\ 4x_1 + 7x_2 - ax_3 &= b \end{aligned}$

3) $\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ ax_1 + 2x_2 + x_3 &= 2 \\ x_1 + ax_2 + ax_3 &= b + 2 \end{aligned}$

4) $\begin{aligned} ax_1 + 2x_2 + 4x_3 &= b + 2 \\ 2x_1 + ax_2 + 4x_3 &= 1 \\ 2x_1 + ax_2 + a^2x_3 &= 1 \end{aligned}$

20. En los sistemas de ecuaciones siguientes, determine los valores de a para los cuales el sistema tiene solución(es). En tal caso determine la(s) solución(es).

1) $\begin{aligned} x_1 + ax_2 &= 1 \\ ax_1 + x_2 &= a \end{aligned}$

2) $\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ ax_1 + x_2 + x_3 &= a \\ a^2x_1 - x_2 + 3x_3 &= a^2 + 1 \end{aligned}$

3) $\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ 2x_1 + 5x_2 + 3x_3 &= 10 \\ 3x_2 + x_3 &= a^2 \end{aligned}$

4) $\begin{aligned} x_1 + 3x_2 + x_3 - x_4 &= 4 \\ 2x_1 + 7x_2 + 4x_3 - 2x_4 &= a^2 + 5 \\ -x_1 - 2x_2 + x_3 + x_4 &= 2 \\ 2x_1 + 8x_2 + 6x_3 - 2x_4 &= 2(a^2 + 1) \end{aligned}$

21. Resuelva los siguientes sistemas homogéneos de ecuaciones lineales:

1) $\begin{aligned} 3x_1 + x_2 + 5x_3 &= 0 \\ 2x_1 + 4x_2 + x_3 &= 0 \\ -4x_1 + 2x_2 - 9x_3 &= 0 \end{aligned}$

2) $\begin{aligned} 2x_1 + x_2 + x_3 &= 0 \\ x_1 + 2x_2 + x_3 &= 0 \\ 2x_1 + 2x_2 + 2x_3 &= 0 \end{aligned}$

3) $\begin{aligned} x_1 + 3x_2 + 3x_3 &= 0 \\ 2x_1 + 7x_2 + 7x_3 &= 0 \\ 2x_1 + 7x_2 + 6x_3 &= 0 \end{aligned}$

4) $\begin{aligned} x_1 + 2x_2 + 2x_3 + x_4 &= 0 \\ x_1 - x_2 + x_3 - 2x_4 &= 0 \\ 2x_1 + x_2 - x_3 + x_4 &= 0 \end{aligned}$

5) $\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 2x_1 - x_2 + 3x_3 &= 0 \end{aligned}$

22. ¿Es verdadero el teorema 2.1 para sistemas no homogéneos de ecuaciones lineales? En caso afirmativo demuéstrelo, de otro modo, dé un contraejemplo.

3. MATRICES (I): OPERACIONES CON MATRICES

El concepto de matriz apareció ya en la sección anterior en relación con los sistemas de ecuaciones lineales, en un intento de ahorro en la notación para el proceso de eliminación Gaussiana. A partir de este momento se va a olvidar esa correspondencia matrices-sistemas de ecuaciones, a menos que se haga referencia explícita a ella. Se emprenderá ahora un estudio de estos nuevos entes matemáticos pensando en ellos como "objetos con existencia propia".

A pesar de que ya fueron establecidas algunas de las definiciones preliminares, éstas se repetirán en su momento para lograr una continuidad en nuestro estudio.

Se llama *matriz* a un arreglo rectangular de números como el siguiente:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

(constituido por m *líneas*, o *filas* o *renglones*, y n *columnas*). A los números a_{ij} ($i = 1, \dots, m, j = 1, \dots, n$) se les llama *elementos* de la matriz.

a_{ij} es, pues, el elemento de la matriz que se encuentra en la i-ésima línea y en la j-ésima columna.

Se usarán letras mayúsculas ($A, B, C, \dots$) para denotar matrices y minúsculas ($a_{ij}, b_{ij}, c_{ij}, \dots$) para denotar a sus elementos.

Se define el *orden* de una matriz como: número de líneas × número de columnas.

Para la matriz A de orden $m \times n$, cuyos elementos son a_{ij}, se usa también la notación $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$.

Si en $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$, ocurre $m = n$, se dice que A es una *matriz cuadrada*. En tal caso se escribe $A = (a_{ij})_{i,j=1,\dots,n}$ y se dice que A es una matriz (cuadrada) de orden n.

Se dice que las matrices A y B son iguales si ellas tienen el mismo orden y sus elementos correspondientes coinciden.

EJEMPLO 1 Por ejemplo, la matriz

$$A = \begin{bmatrix} 1 & 2 \\ 8 & 3 \end{bmatrix}$$

es igual a la matriz

$$B = \begin{bmatrix} 1 & \gamma \\ 8 & 3 \end{bmatrix}$$

si y sólo si, $\gamma = 2$, mientras que la matriz

$$C = \begin{bmatrix} 1 & 2 & 0 \\ 8 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

es *diferente* a la matriz A (sus órdenes no coinciden).

Dada la matriz cuadrada A de orden n, $A = (a_{ij})_{i,j=1,\dots,n}$, se dice que los elementos a_{ii}, $i = 1, \dots, n$, constituyen la *diagonal principal* de la matriz.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

A la matriz de orden n

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

en la cual todos sus elementos son cero, excepto los de la diagonal principal, que son uno, se le llama *matriz identidad* (de orden n), y se denota por I_n. Así, por ejemplo:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Se define también la matriz cero de orden $m \times n$, denotada por $0_{m \times n}$(si el orden es claro en el contexto, se va a escribir sólo 0), como la matriz que tiene todos sus elementos iguales a cero. Por ejemplo:

$$0_{2 \times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad 0_{3 \times 4} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Pásese ahora a ver qué tipo de operaciones se pueden realizar con las matrices y cómo se efectúan.

3.1. SUMA Y PRODUCTO POR UN ESCALAR

DEFINICIÓN 3.1 (Suma de matrices)

Sean las matrices $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$, y $B = (b_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$. Se define la suma de A y B, denotada por $A + B$, como la matriz $(c_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ en donde $c_{ij} = a_{ij} + b_{ij}$. Es decir,

$$A + B = (a_{ij} + b_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

O bien

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ & & \dots & \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ & & \dots & \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix}$$

Véase que la suma de matrices se ha definido solamente para matrices del mismo orden.

EJEMPLO 2

Por ejemplo:

$$\begin{bmatrix} 1 & 2 & -1 \\ 3 & 2 & 4 \\ 5 & 7 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 3 & 3 \\ -2 & 5 & 2 \\ 7 & -1 & 5 \end{bmatrix} = \begin{bmatrix} 5 & 5 & 2 \\ 1 & 7 & 6 \\ 12 & 6 & 7 \end{bmatrix}$$

*DEFINICIÓN 3.2 (Producto por un escalar)**

Sea la matriz $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$, y sea k un escalar (dígase $k \in R$).

Se define el producto de la matriz A por el escalar k, denotado por kA, como la matriz $(c_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$, en donde $c_{ij} = ka_{ij}$.

Es decir,

$$kA = (ka_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

o bien

$$k \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ & & \dots & \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{bmatrix}$$

EJEMPLO 3

Por ejemplo

$$3 \begin{bmatrix} 2 & 1 & 4 & 3 \\ -1 & 2 & 5 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 3 & 12 & 9 \\ -3 & 6 & 15 & 3 \end{bmatrix}$$

Aunque estas operaciones no parezcan a primera vista muy emocionantes, resulta (como se aclarará en el capítulo 3) que el conjunto (inerme) de matrices se convierte, con las operaciones anteriormente definidas, en una estructura algebraica muy importante: la de espacio vectorial.

---

*El término *escalar* y su uso para denotar esta operación se hará claro en el capítulo 3. Por el momento, identifíquese *escalar* con "número" (real o complejo).

Las principales propiedades de que gozan estas operaciones están contenidas en el siguiente teorema:

TEOREMA 3.1

> Sean A, B y C tres matrices cualesquiera del mismo orden ($m \times n$) con elementos a_{ij}, b_{ij} y c_{ij} respectivamente. Sean k y l dos escalares.
> Entonces:
>
> 1) $A + B = B + A$ (la suma es conmutativa).
> 2) $A + (B + C) = (A + B) + C$ (la suma es asociativa)
> 3) Existe una matriz "cero", tal que $A + 0 = A$.
> 4) Dada la matriz A, existe una matriz $(-A)$ tal que $A + (-A) = 0$ (la matriz cero).
> 5) $k(A + B) = kA + kB$
> 6) $(k + l)A = ka + la$
> 7) $(kl)A = k(lA)$
> 8) $1 \cdot A = A$

DEMOSTRACIÓN Se trata de verificaciones de simple rutina, en las que se emplean propiedades conocidas de los números reales. Por ejemplo, $A + B$ es una matriz $m \times n$ cuyos elementos son $a_{ij} + b_{ij}$, mientras que $B + A$ es una matriz $m \times n$ cuyos elementos son $b_{ij} + a_{ij}$. Son, pues, matrices del mismo orden y además, por la propiedad conmutativa de la suma de números reales, $a_{ij} + b_{ij} = b_{ij} + a_{ij}$, es decir, sus elementos correspondientes coinciden. Entonces $A + B = B + A$. El resto se deja de ejercicio al lector.

Q.E.D.

3.2 PRODUCTO DE MATRICES

Además de las operaciones descritas en la subsección (3.1), con las matrices se puede efectuar otra operación sumamente interesante: se pueden multiplicar matrices entre sí (bajo ciertas condiciones, como se verá a continuación).

La definición que se dará de multiplicación de matrices es, sin duda, la *menos natural* posible. Sin embargo, a pesar de que resultará muy sofisticada la manera de multiplicar una matriz con otra, será también, sin duda, la manera *más conveniente* de hacerlo.

DEFINICIÓN 3.3 Sea A una matriz de orden $m \times p$ con elementos a_{ij}, y B una matriz de orden $p \times n$ con elementos b_{ij}. Se define el *producto* de A por B, denotado por AB, como la matriz de orden $m \times n$ con elementos c_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ definidos por c_{ij}

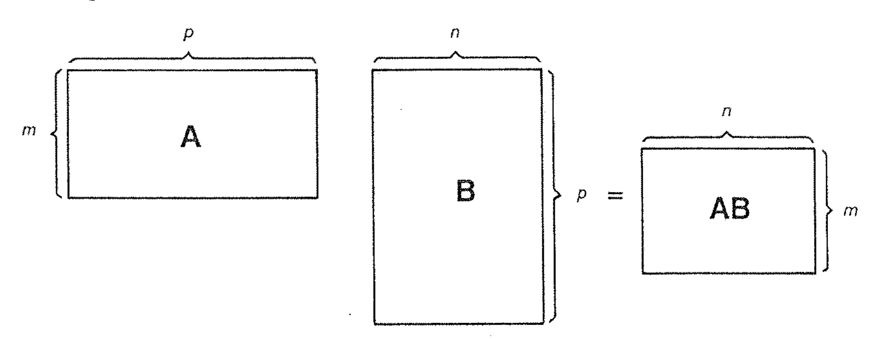
$$C_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$$

Es decir

$$AB = \left(\sum_{k=1}^{p} a_{ik}b_{kj}\right)_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

Obsérvese que el producto de A por B se ha definido sólo en el caso en que el número de columnas de A coincide con el número de líneas de B, en cuyo caso la matriz producto tendrá orden (número de líneas de A) $\times$ (número de columnas de B).

Esquemáticamente



La manera más fácil de aprender cómo se efectúa la operación de producto entre matrices es simplemente recordar la frase "líneas (de A) por columnas (de B)".

Más concretamente, el elemento de la i-ésima línea y j-ésima columna de AB se obtiene "multiplicando" la i-ésima línea de A con la j-ésima columna de B (o sea, multiplicando el r-ésimo elemento de la i-ésima línea de A con el r-ésimo elemento de la j-ésima columna de B, donde $r = 1, \ldots, p$ = número de columnas de A = número de líneas de B, y luego sumando estos productos).

Esquemáticamente

columna j columna j

línea i $\begin{bmatrix} & & & \\ a_{i1} & a_{i2} & \ldots & a_{ip} \\ & & & \end{bmatrix} \begin{bmatrix} & b_{1j} & \\ & b_{2j} & \\ & \vdots & \\ & b_{pj} & \end{bmatrix} = \begin{bmatrix} & & \\ & c_{ij} & \\ & & \end{bmatrix}$ línea i

A B AB

donde

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$$

EJEMPLO 4 Por ejemplo

$$\begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix}\begin{bmatrix} 5 & -2 \\ 6 & 8 \end{bmatrix} = \begin{bmatrix} (2)(5)+(3)(6) & (2)(-2)+(3)(8) \\ (4)(5)+(-1)(6) & (4)(-2)+(-1)(8) \end{bmatrix} = \begin{bmatrix} 28 & 20 \\ 14 & -16 \end{bmatrix}$$

Existen razones de fondo para definir el producto de matrices de la manera como se hizo. La más importante podrá verse en el capítulo 4 (aparecerá en relación al estudio de la representación matricial de una composición de transformaciones lineales entre espacios vectoriales de dimensión finita). Sin embargo, en este

momento se puede ver un problema relacionado con sistemas de ecuaciones en donde aparece naturalmente la multiplicación de matrices. *

Supóngase que se tiene el sistema de 3 ecuaciones con 4 incógnitas.

$$\begin{aligned} x_1 + 2x_2 + 6x_3 - x_4 &= 0 \\ 2x_1 - x_2 + x_3 + x_4 &= 0 \qquad (A) \\ 3x_1 - 2x_2 + 2x_3 - 3x_4 &= 0 \end{aligned}$$

Se quieren cambiar las variables x_1, x_2, x_3 y x_4 de este sistema por las nuevas variables y_1, y_2 relacionadas por

$$\begin{aligned} x_1 &= 2y_1 + y_2 \\ x_2 &= -y_1 + 3y_2 \\ x_3 &= y_1 + y_2 \qquad (B) \\ x_4 &= 3y_1 - 2y_2 \end{aligned}$$

Al hacer las sustituciones requeridas se obtiene

$$\begin{aligned} (2y_1 + y_2) + 2(-y_1 + 3y_2) + 6(y_1 + y_2) - (3y_1 - 2y_2) &= 0 \\ 2(2y_1 + y_2) - (-y_1 + 3y_2) + (y_1 + y_2) + (3y_1 - 2y_2) &= 0 \\ 3(2y_1 + y_2) - 2(-y_1 + 3y_2) + 2(y_1 + y_2) - 3(3y_1 - 2y_2) &= 0 \end{aligned}$$

o sea

$$\begin{aligned} 3y_1 + 15y_2 &= 0 \\ 9y_1 - 2y_2 &= 0 \qquad (C) \\ y_1 + 5y_2 &= 0 \end{aligned}$$

Obsérvese que la matriz del sistema (A) es:

$$A = \begin{bmatrix} 1 & 2 & 6 & -1 \\ 2 & -1 & 1 & 1 \\ 3 & -2 & 2 & -3 \end{bmatrix}$$

Las ecuaciones de cambio de variable (B) pueden también ser contempladas como un sistema de ecuaciones, en donde x_1, x_2, x_3, x_4 son dados y las incógnitas son y_1, y_2, La matriz de este sistema es:

$$B = \begin{bmatrix} 2 & 1 \\ -1 & 3 \\ 1 & 1 \\ 3 & -2 \end{bmatrix}$$

*Comentario para ser leído después de estudiar el capítulo 4: En realidad éste es un problema de composición de transformaciones lineales: se está realizando la sustitución de un sistema lineal en otro.

Asimismo, la matriz del sistema (C) en las nuevas incógnitas y_1, y_2 es:

$$C = \begin{bmatrix} 3 & 15 \\ 9 & -2 \\ 1 & 5 \end{bmatrix}$$

Pero

$$AB = \begin{bmatrix} 1 & 2 & 6 & -1 \\ 2 & -1 & 1 & 1 \\ 3 & -2 & 2 & -3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 3 \\ 1 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 3 & 15 \\ 9 & -2 \\ 1 & 5 \end{bmatrix} = C$$

Es decir, el sistema "compuesto" [el resultante de la sustitución de (B) en (A)] tiene por matriz el producto de las matrices de (A) y de (B).

También relacionado con sistemas de ecuaciones, se puede ver que el producto de matrices tiene una gran ventaja para denotar "matricialmente" un sistema.

Considérese el sistema de m ecuaciones con n incógnitas

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

Sea $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ la matriz $(m \times n)$ del sistema. Considérese también la matriz X de orden $n \times 1$ formada por las incógnitas $x_1, x_2, \ldots, x_n$ y la matriz B de orden $m \times 1$ formada por los términos independientes $b_1, b_2, \ldots, b_m$.

Entonces el sistema puede escribirse como

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & \ldots & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

o sea

$$AX = B$$

(donde AX representa la multiplicación de la matriz A por la matriz X y el signo = indica una igualdad entre matrices).

EJEMPLO 5

Por ejemplo, el sistema

$$\begin{aligned} 3x_1 + 2x_2 - 6x_3 &= 1 \\ x_1 - x_2 + 2x_3 &= 2 \end{aligned}$$

se denota matricialmente como $AX = B$, en donde

$$A = \begin{bmatrix} 3 & 2 & -6 \\ 1 & -1 & 2 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Las propiedades más importantes de que goza la multiplicación de matrices se encuentra en el siguiente teorema:

TEOREMA 3.2

Sean A, B y C tres matrices de órdenes tales que las operaciones indicadas en seguida tienen sentido. Sea k un escalar. Entonces

1) $A(BC) = (AB)C$ (la multiplicación es asociativa).
2) $A(B + C) = AB + AC$ (distributividad).
3) $(A + B)C = AC + BC$
4) $k(AB) = (kA)B = A(kB)$
5) $IA = A$, $AI = A$, donde I es la matriz identidad.
6) $0 \cdot A = 0$, $A \cdot 0 = 0$, donde 0 es la matriz cero.

DEMOSTRACIÓN Se hará la demostración de la primera propiedad (asociatividad), dejando como ejercicio al lector las restantes. Sea A una matriz $m \times p$ de elementos a_{ij}, B una matriz $p \times r$ de elementos b_{ij}, y C una matriz $r \times n$ de elementos c_{ij}. Tienen pues sentido todos los productos indicados. Se ve fácilmente que tanto la matriz $A(BC)$ como la matriz $(AB)C$ tienen orden $m \times n$. Entonces sólo se debe de verificar que sus elementos correspondientes coinciden. Llámese θ_{ij} (ρ_{ij}) al elemento de la i-ésima línea y j-ésima columna de $A(BC)$ (de $(AB)C$ respectivamente). Entonces

$$\theta_{ij} = \sum_{k=1}^{p} a_{ik}\,\beta_{kj}$$

donde β_{kj} es el elemento de la k-ésima línea y j-ésima columna de BC, o sea

$$\beta_{kj} = \sum_{s=1}^{r} b_{ks}\,c_{sj}$$

Entonces

$$\theta_{ij} = \sum_{k=1}^{p} a_{ik} \sum_{s=1}^{r} b_{ks}\,c_{sj} = \sum_{k=1}^{p} \sum_{s=1}^{r} a_{ik}\,b_{ks}\,c_{sj}$$

Por otra parte

$$\rho_{ij} = \sum_{s=1}^{r} \alpha_{is}\,c_{sj}$$

en donde α_{is} es el elemento de la i-ésima línea y s-ésima columna de AB, o sea

$$\alpha_{is} = \sum_{k=1}^{p} a_{ik}\,b_{ks}$$

Entonces

$$\rho_{ij} = \sum_{s=1}^{r} \alpha_{is}c_{sj} = \sum_{s=1}^{r} \sum_{k=1}^{p} a_{ik}\,b_{ks}\,c_{sj} = \theta_{ij}$$

Q.E.D.

Se puede ver entonces que la multiplicación de matrices goza de muchas propiedades análogas a la multiplicación de números reales. Es asociativa, es distributiva, la matriz identidad I actúa neutramente en el producto (como el 1 en los números reales), etc. Sin embargo, existen algunas propiedades del producto de números reales, que no son válidas para matrices.

Primeramente obsérvese que, dadas las matrices A y B para las cuales tiene sentido el producto AB, el producto BA podría no tener sentido (por ejemplo, si A es de orden $m \times r$ y B de orden $r \times p$ con $m \neq p$). Aún más, si A y B fueran matrices cuadradas, en cuyo caso se pueden realizar los productos AB y BA, no es cierto en general que $AB = BA$. En realidad, la propiedad $AB = BA$ es una propiedad muy especial; en este caso se dice que A y B son matrices que conmutan. La multiplicación entre matrices, es pues, *no conmutativa*.

EJEMPLO 6 Por ejemplo, si A es una matriz 3×2 y B una matriz 2×3 AB y BA no pueden ser iguales, pues sus órdenes son distintos (AB de orden 3 y BA de orden 2). Aún si AB y BA tienen el mismo orden, puede ocurrir $AB \neq BA$, por ejemplo:

$$A = \begin{bmatrix} 1 & 2 \\ 5 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 3 \\ 2 & 1 \end{bmatrix}, \quad AB = \begin{bmatrix} 3 & 5 \\ -7 & 14 \end{bmatrix} \neq \begin{bmatrix} 14 & -5 \\ 7 & 3 \end{bmatrix} = BA$$

Por otro lado, tampoco es válida la "ley de la cancelación"* para matrices: si

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & -2 \\ 3 & 10 \end{bmatrix}$$

se tiene

$$AB = \begin{bmatrix} 12 & 16 \\ 6 & 8 \end{bmatrix} = AC$$

pero $B \neq C$.

*Si $a, b, c \in R$ y $ab = ac$, $a \neq 0$, entonces $b = c$.

Por último, la propiedad de los números reales "$a, b \in R$, $ab = 0 \Rightarrow a = 0$ o $b = 0$", también es falsa para matrices. Por ejemplo:

$$\begin{bmatrix} 2 & -1 \\ 8 & -4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

pero ambas matrices en el producto son distintas de (la matriz) cero.

3.3 PARTICIÓN DE MATRICES

Considérese la matriz $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ de orden $m \times n$. En el arreglo rectangular de números que representa A, se dibujan líneas verticales y horizontales como se muestra a continuación:

$$A = \left[\begin{array}{c|cc|c|cc} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ \hline a_{31} & a_{32} & a_{33} & \cdots & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \hline a_{m1} & a_{m2} & a_{m3} & \cdots & a_{m,n-1} & a_{mn} \end{array}\right]$$

La matriz A queda entonces dividida en bloques rectangulares de sus elementos. Se dice así que se ha efectuado una *partición* en la matriz A. Los bloques resultantes de esta partición pueden ser considerados por sí mismos como matrices; se llamarán *submatrices* de la matriz A.

EJEMPLO 7 Por ejemplo, si A es la matriz

$$A = \begin{bmatrix} 3 & 1 & 4 & 0 \\ 2 & 1 & 3 & 2 \\ -1 & 3 & 4 & 5 \end{bmatrix}$$

se puede efectuar en ella la partición

$$A = \left[\begin{array}{ccc|c} 3 & 1 & 4 & 0 \\ 2 & 1 & 3 & 2 \\ \hline -1 & 3 & 4 & 5 \end{array}\right]$$

En tal caso se escribe

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

en donde A_{ij}, $i, j = 1, 2$ son las submatrices resultantes de la partición de A

$$A_{11} = \begin{bmatrix} 3 & 1 & 4 \\ 2 & 1 & 3 \end{bmatrix}, \; A_{12} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \; A_{21} = [\,-1 \quad 3 \quad 4\,], \; A_{22} = [5]$$

En la matriz A anterior se puede efectuar también la partición

$$A = \left[\begin{array}{c|cc|c} 3 & 1 & 4 & 0 \\ \hline 2 & 1 & 3 & 2 \\ -1 & 3 & 4 & 5 \end{array}\right]$$

Entonces se escribe

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

en donde

$$A_{11} = [3] \quad A_{12} = [\,1 \quad 4\,] \; A_{13} = [0]$$

$$A_{21} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} A_{22} = \begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix} A_{23} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

Es claro que existen muchas maneras de efectuar particiones en una matriz. Fundamentalmente existen 3 razones por las que puede ser importante efectuar una determinada partición en una matriz:

1) Para simplificar la notación. Por ejemplo, supóngase que se tiene la matriz

$$A = \begin{bmatrix} 1 & 0 & 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Esta misma matriz se puede escribir como:

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

en donde

$$A_{11}=\begin{bmatrix}1&0&2\\0&2&1\\0&0&3\end{bmatrix}\qquad A_{22}=\begin{bmatrix}3&1&4\\0&1&1\\0&0&1\end{bmatrix}$$

y los ceros que aparecen en la forma simplificada de la matriz A se refieren a la matriz cero.

2) Para resaltar algún aspecto de la estructura de la matriz. Por ejemplo, si A es la matriz

$$A=\begin{bmatrix}2&8&20&1&0&0\\-4&0&5&0&1&0\\3&0&0&0&0&1\\1&0&0&0&7&0\\0&1&0&3&5&0\\0&0&1&1&1&0\end{bmatrix}$$

se puede escribir

$$A=\begin{bmatrix}A_{11}&I_3\\I_3&A_{22}\end{bmatrix}$$

en donde

$$A_{11}=\begin{bmatrix}2&8&20\\-4&0&5\\3&0&0\end{bmatrix}\qquad A_{22}=\begin{bmatrix}0&7&0\\3&5&0\\1&1&0\end{bmatrix}$$

quedando así explícita la existencia de dos submatrices identidad en la matriz A.

3) Para simplificar los cálculos con matrices, especialmente el producto de ellas.

Éste es el punto que se discutirá con cierto detalle. Se verá cómo se puede efectuar la multiplicación de dos matrices en las que se ha realizado previamente una determinada partición.

Se concentrará el estudio, con fines de simplificación, al caso en el que las matrices correspondientes quedan divididas en cuatro submatrices.

Sea A una matriz de orden $m \times p$ y B una matriz de orden $p \times n$. Dígase que en la matriz A se ha efectuado la partición que permite escribirla como

$$A=\left[\begin{array}{c|c}\overbrace{A_{11}}^{p_1}&\overbrace{A_{12}}^{p_2}\\\hline A_{21}&A_{22}\end{array}\right]\begin{array}{l}\}\ m_1\\\}\ m_2\end{array}$$

en donde

$$p_1+p_2=p\quad,\quad m_1+m_2=m$$

y que la matriz B queda escrita como

$$B=\left[\begin{array}{c|c}\overbrace{B_{11}}^{n_1}&\overbrace{B_{12}}^{n_2}\\\hline B_{21}&B_{22}\end{array}\right]\begin{array}{l}\}\ p_1\\\}\ p_2\end{array}$$

en donde

$$n_1+n_2=n$$

Véase que del mismo modo como se efectuó la partición en las columnas de A, se realizó en las líneas de B.

Se demostrará enseguida que el producto AB puede efectuarse como

$$A=\begin{bmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{bmatrix}\begin{bmatrix}B_{11}&B_{12}\\B_{21}&B_{22}\end{bmatrix}=\begin{bmatrix}A_{11}B_{11}+A_{12}B_{21}&A_{11}B_{12}+A_{12}B_{22}\\A_{21}B_{11}+A_{22}B_{21}&A_{21}B_{12}+A_{22}B_{22}\end{bmatrix}$$

Es decir, que la multiplicación entre dos matrices A y B en las que se ha efectuado una partición (de modo que la división que se hizo en las columnas de A coincide con la hecha en las líneas de B) puede efectuarse como si las submatrices resultantes de las particiones correspondientes fueran elementos de las matrices A y B, y realizando entonces la multiplicación entre ellas según lo discutido en la subsección anterior.

Si se escribe el producto AB como

$$AB=\left[\begin{array}{c|c}\overbrace{C_{11}}^{n_1}&\overbrace{C_{12}}^{n_2}\\\hline C_{21}&C_{22}\end{array}\right]\begin{array}{l}\}\ m_1\\\}\ m_2\end{array}$$

se debe entonces mostrar que

$$C_{11}=A_{11}B_{11}+A_{12}B_{21}\qquad C_{21}=A_{21}B_{11}+A_{22}B_{21}$$
$$C_{12}=A_{11}B_{12}+A_{12}B_{22}\qquad C_{22}=A_{21}B_{12}+A_{22}B_{22}$$

Obsérvese primeramente que todas las operaciones entre matrices indicadas en las fórmulas anteriores están bien definidas. Se demostrará solamente la primera de estas fórmulas siendo la demostración de las restantes completamente análoga a la que se hará y quedando como ejercicio para el lector. Se tiene entonces que el producto de la matriz A_{11} de orden $m_1 \times p_1$ por la matriz B_{11} de orden $p_1 \times n_1$, es la matriz $A_{11}B_{11}$ de orden $m_1 \times n_1$ dada por

$$A_{11}B_{11} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p_1} \\ a_{21} & a_{22} & \cdots & a_{2p_1} \\ & \cdots & & \\ a_{m_11} & a_{m_12} & \cdots & a_{m_1p_1} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n_1} \\ b_{21} & b_{22} & \cdots & b_{2n_1} \\ & \cdots & & \\ b_{m_11} & b_{p_12} & \cdots & b_{p_1n_1} \end{bmatrix} = \left[\begin{array}{c:c} & \\ \cdots & \alpha_{ij} \cdots \\ & \end{array}\right]$$

en donde

$$\alpha_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip_1}b_{p_1j} \qquad \begin{matrix} 1 \le i \le m_1 \\ 1 \le j \le n_1 \end{matrix}$$

Similarmente, el producto de la matriz A_{12} de orden $m_1 \times p_2$ por la matriz B_{21} de orden $p_2 \times n_1$, es la matriz $A_{12}B_{21}$ de orden $m_1 \times n_1$ dada por

$$A_{12}B_{21} = \begin{bmatrix} a_{1,p_1+1} & a_{1,p_1+2} & \cdots & a_{1p} \\ a_{2,p_1+1} & a_{2,p_1+2} & \cdots & a_{2p} \\ & \cdots & & \\ a_{m_1,p_1+1} & a_{m_1,p_1+2} & \cdots & a_{m_1p} \end{bmatrix} \begin{bmatrix} b_{p_1+1,1} & b_{p_1+1,2} & \cdots & b_{p_1+1,n_1} \\ b_{p_1+2,1} & b_{p_1+2,2} & \cdots & b_{p_1+2,n_1} \\ & \cdots & & \\ b_{p_1} & b_{p_2} & \cdots & b_{pn_1} \end{bmatrix} = \left[\begin{array}{c:c} & \\ \cdots & \beta_{ij} \cdots \\ & \end{array}\right]$$

en donde

$$\beta_{ij} = a_{i,p_1+1}b_{p_1+1,j} + a_{i,p_1+2}b_{p_1+2,j} + \ldots + a_{ip}b_{pj}, \qquad \begin{matrix} 1 \le i \le m_1 \\ 1 \le j \le n_1 \end{matrix}$$

Entonces la matriz $C_{11} = A_{11}B_{11} + A_{12}B_{21}$ tiene en su i-ésima línea y j-ésima columna al elemento $\alpha_{ij} + \beta_{ij}$ el cual es

$$a_{i1}b_{1j} + \ldots + a_{ip_1}b_{p_1j} + a_{i,p_1+1}b_{p_1+1,j} + \ldots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj} \qquad \begin{matrix} 1 \le i \le m_1 \\ 1 \le j \le n_1 \end{matrix}$$

Éste es precisamente el elemento de la i-ésima línea y j-ésima columna ($1 \le i \le m_1$, $1 \le j \le n_1$) del producto AB. Se ha pues mostrado que

$$AB = \left[\begin{array}{c:c} \overbrace{A_{11}B_{11}+A_{12}B_{21}}^{n_1} & \\ \hdashline & \\ \end{array}\right] \} \; m_1$$

El resto, como ya se había dicho, queda como ejercicio para el lector.

EJEMPLO 8 Por ejemplo, sean A y B las matrices

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -1 & 4 \\ 3 & 1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 3 & 4 \\ -2 & 2 & 3 \end{bmatrix}$$

La matriz AB es:

$$AB = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -1 & 4 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 2 & 3 & 4 \\ -2 & 2 & 3 \end{bmatrix} = \begin{bmatrix} -2 & 11 & 11 \\ -10 & 5 & 8 \\ 3 & 8 & 4 \end{bmatrix}$$

Este mismo resultado se puede obtener haciendo una previa partición de las matrices A y B y procediendo según lo analizado anteriormente. Por ejemplo:

$$AB = \left[\begin{array}{c:cc} 2 & 1 & 3 \\ 0 & -1 & 4 \\ \hdashline 3 & 1 & 1 \end{array}\right] \left[\begin{array}{c:cc} 1 & 1 & -1 \\ \hdashline 2 & 3 & 4 \\ -2 & 2 & 3 \end{array}\right] = \left[\begin{array}{c:c} \begin{bmatrix} 2 \\ 0 \end{bmatrix}[1] + \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ -2 \end{bmatrix} & \begin{bmatrix} 2 \\ 0 \end{bmatrix}[1 \;\; -1] + \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix} \\ \hdashline [3][1] + [1 \;\; 1]\begin{bmatrix} 2 \\ -2 \end{bmatrix} & [3][1 \;\; -1] + [1 \;\; 1]\begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix} \end{array}\right]$$

$$\left[\begin{array}{c:c} \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \begin{bmatrix} -4 \\ -10 \end{bmatrix} & \begin{bmatrix} 2 & -2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 9 & 13 \\ 5 & 8 \end{bmatrix} \\ \hdashline [3] + [0] & [3 \;\; -3] + [5 \;\; 7] \end{array}\right] = \left[\begin{array}{c:c} \begin{bmatrix} -2 \\ -10 \end{bmatrix} & \begin{bmatrix} 11 & 11 \\ 5 & 8 \end{bmatrix} \\ \hdashline [3] & [8 \;\; 4] \end{array}\right]$$

resultado que coincide (después de quitar los corchetes interiores) con el obtenido directamente para el producto AB. Se hubiera podido efectuar una partición distinta en las matrices A y B y obtener el mismo resultado. El lector puede verificar, por ejemplo, que

$$AB = \left[\begin{array}{cc:c} 2 & 1 & 3 \\ 0 & -1 & 4 \\ \hdashline 3 & 1 & 1 \end{array}\right] \left[\begin{array}{cc:c} 1 & 1 & -1 \\ 2 & 3 & 4 \\ \hdashline -2 & 2 & 3 \end{array}\right] = \left[\begin{array}{cc:c} -2 & 11 & 11 \\ -10 & 5 & 8 \\ \hdashline 3 & 8 & 4 \end{array}\right]$$

EJEMPLO 9 Como un último ejemplo, considere las matrices

$$A = \begin{bmatrix} 1 & 3 & -1 & 0 & 0 & 0 \\ 2 & 1 & 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 1 & 0 \\ 1 & 3 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & 1 & 3 \end{bmatrix}$$

Supóngase que se quiere calcular el producto AB. Se verá que con una partición adecuada de estas matrices, este producto es muy fácil de calcular.

Sean

$$A_1 = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 2 \\ 3 & 4 & 0 \end{bmatrix} \quad B_1 = \begin{bmatrix} 1 & 0 & 1 \\ -1 & -1 & 2 \\ 1 & 3 & 1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

Entonces A y B se ven como

$$A = \begin{bmatrix} A_1 & 0 \\ I & I \end{bmatrix} \quad B = \begin{bmatrix} B_1 & I \\ 0 & B_2 \end{bmatrix}$$

en donde 0 e I son las matrices cero e identidad de orden 3, respectivamente. Por lo tanto

$$AB = \begin{bmatrix} A_1 & 0 \\ I & I \end{bmatrix} \begin{bmatrix} B_1 & I \\ 0 & B_2 \end{bmatrix} = \begin{bmatrix} A_1B_1 + 0 \cdot 0 & A_1 + 0 \cdot B_2 \\ I \cdot B_1 + I \cdot 0 & I \cdot I + I \cdot B_2 \end{bmatrix} = \begin{bmatrix} A_1B_1 & A_1 \\ B_1 & I + B_2 \end{bmatrix}$$

Todo lo que se tiene que hacer es calcular el producto A_1B_1

$$A_1B_1 = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 2 \\ 3 & 4 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -1 & -1 & 2 \\ 1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} -3 & -6 & 6 \\ 3 & 5 & 6 \\ -1 & -4 & 11 \end{bmatrix}$$

Entonces

$$AB = \begin{bmatrix} A_1B_1 & A_1 \\ B_1 & I + B_2 \end{bmatrix} = \begin{bmatrix} -3 & -6 & 6 & 1 & 3 & -1 \\ 3 & 5 & 6 & 2 & 1 & 2 \\ -1 & -4 & 11 & 3 & 4 & 0 \\ 1 & 0 & 1 & 3 & 1 & 1 \\ -1 & -1 & 2 & 1 & 2 & 1 \\ 1 & 3 & 1 & 2 & 1 & 4 \end{bmatrix}$$

3.4. LA SOLUCIÓN DE UN SISTEMA NO HOMOGÉNEO DE ECUACIONES LINEALES

Considérese el sistema de m ecuaciones con n incógnitas $AX = B$ (A es entonces una matriz $m \times n$). Se asociará a éste el sistema homogéneo $AX = 0$ (aquí, 0 denota la matriz cero $m \times 1$) y se dirá que este sistema es el sistema homogéneo asociado a $AX = B$.

Se dirá que

$$\tilde{X} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_n \end{bmatrix}$$

es una solución de $AX = B$ si la matriz $A\tilde{X}$ es idéntica a la matriz B (se trata del mismo concepto de solución de ecuaciones lineales, escrito matricialmente), lo cual se escribirá simplemente como: si $\tilde{X}$ es solución de $AX = B$ entonces $A\tilde{X} = B$ (distinga entre *el sistema* $AX = B$ y la igualdad de matrices-numéricas $A\tilde{X} = B$).

Supóngase que el sistema $AX = B$ es consistente (su conjunto solución no es vacío). Sea $\tilde{X}_p$ una solución particular del sistema, y sea $\tilde{X}$ *cualquier* otra solución del mismo.

Al usarse las propiedades de la multiplicación matricial descritas en el teorema (3.2) se puede escribir entonces

$$A(\tilde{X} - \tilde{X}_p) = A\tilde{X} - A\tilde{X}_p = B - B = 0$$

Entonces $\tilde{X}_h = \tilde{X} - \tilde{X}_p$ es alguna solución del sistema homogéneo asociado $AX = 0$. Es decir, como $\tilde{X} = \tilde{X}_p + \tilde{X}_h$, cualquier solución del sistema $AX = B$ se escribe como la suma de una solución particular $\tilde{X}_p$ del mismo, más alguna solución del sistema homogéneo asociado.

Sea ahora $\tilde{X}_h$ *cualquier* solución de $AX = 0$. Al escribir $\tilde{X} = \tilde{X}_p + \tilde{X}_h$, se ve que $A\tilde{X} = B$, esto es, $\tilde{X}$ es solución de $AX = B$

Entonces, si $\tilde{X}$ es *cualquier* solución de $AX = B$, ésta puede escribirse como la suma de una solución particular $\tilde{X}_p$ del sistema más $\tilde{X}_h$, en donde $\tilde{X}_h$ es *cualquier* solución del sistema homogéneo asociado.

La manera correcta de referirse a $\tilde{X}$ como "cualquier" solución de $AX = B$ (o a $\tilde{X}_h$ de $AX = 0$) es diciendo que $\tilde{X}$ (respectivamente $\tilde{X}_h$) es la *solución general* del sistema $AX = B$ (respectivamente, del sistema $AX = 0$).

Se ha entonces probado el siguiente resultado.

TEOREMA 3.3

Si el sistema $AX = B$ es consistente, entonces la solución general del sistema puede escribirse como la suma de una solución particular del mismo más la solución general del sistema homogéneo asociado $AX = 0$.

COROLARIO

El sistema de ecuaciones $AX = B$
1) No tiene solución,
2) tiene una única solución, o bien,
3) tiene una infinidad de soluciones.

DEMOSTRACIÓN Primeramente obsérvese que el sistema homogéneo $AX = 0$, que siempre es consistente, de tener otras soluciones además de la trivial, tendrá una infinidad de ellas. En efecto, si X_h es una solución no trivial de $AX = 0$, entonces tX_h, $t \in R$ será también solución, pues $A(tX_h) = t(AX_h) = t0 = 0$. Supóngase entonces que $AX = B$ es consistente. Sea X la solución general del sistema. Entonces $X = X_p + X_h$, como en el teorema anterior. Si el sistema $AX = 0$ tiene sólo la solución trivial (X_h es la matriz cero), entonces $X = X_p$ y la solución del sistema es única (véase que la

afirmación recíproca también es cierta). Caso contrario (es decir, $AX = 0$ tiene una infinidad de soluciones), es claro que el sistema tiene un número infinito de soluciones.

Q.E.D.

EJEMPLO 10 Para ilustrar el resultado al que se refiere el teorema anterior, considere el sistema

$$\begin{aligned} x_1 + 6x_2 - 5x_3 &= 1 \\ x_1 + 2x_2 - x_3 &= 5 \\ 3x_1 + 2x_2 + x_3 &= 19 \end{aligned}$$

Al usar eliminación Gaussiana, se obtiene

$$\begin{bmatrix} 1 & 6 & -5 & 1 \\ 1 & 2 & -1 & 5 \\ 3 & 2 & 1 & 19 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 1 & 7 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

por lo que la solución general del sistema es:

$$\begin{aligned} x_1 &= 7 - t \\ x_2 &= -1 + t \qquad (t \in R) \\ x_3 &= t \end{aligned}$$

o bien, en notación matricial

$$\tilde{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 - t \\ -1 + t \\ t \end{bmatrix}$$

Obsérvese que

$$\tilde{X} = \begin{bmatrix} 7 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} -t \\ t \\ t \end{bmatrix}$$

en donde $x_1 = 7$, $x_2 = -1$, $x_3 = 0$ es una solución particular del sistema y $x_1 = -t$, $x_2 = t$, $x_3 = t$ ($t \in R$) es la solución general del sistema homogéneo asociado.

EJERCICIOS (SECCIÓN 3, CAPÍTULO 1)

1. Determine los valores de a y b para que las matrices

$$\begin{bmatrix} 2 & 3 \\ 5 & 4 \end{bmatrix} \text{ y } \begin{bmatrix} 2a + b & 3 \\ a - b & 4 \end{bmatrix}$$

sean iguales

2. Determine los valores de a y b de modo que se tenga la igualdad

$$a\begin{bmatrix} 2 \\ -1 \end{bmatrix} - b\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

3. Determine los valores de a, b, c y d de tal modo que se obtenga la igualdad

$$\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

4. Dadas las matrices A y B y el número a, calcule en cada inciso $A + B$, aA, aB, $a(A+B)$, $aA + aB$ y compruebe en cada caso la validez del inciso (5) del teorema 3.1.

1) $A = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, $a = 4$

2) $A = \begin{bmatrix} 2 & 5 & 4 \\ -2 & 3 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 2 & -1 \\ 3 & 2 & 2 \end{bmatrix}$, $a = 8$

3) $A = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 2 & -2 \\ 4 & 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 5 & 4 \\ 3 & 2 & 4 \end{bmatrix}$, $a = -2$

5. Dadas las matrices A y B y el número k, calcule en cada inciso AB, kA, kB, $k(AB)$, $(kA)(B)$, $A(kB)$ y demuestre, en cada, caso la validez del inciso (4) del teorema 3.2.

a) $A = \begin{bmatrix} 2 & 3 & 1 \\ 2 & 2 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 5 \\ 2 & 1 \\ 3 & 3 \end{bmatrix}$, $k = 2$

b) $A = \begin{bmatrix} 0 & -1 \\ 1 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 4 \\ 8 & 5 \end{bmatrix}$, $k = -4$

c) $A = [1 \quad 3 \quad 4]$, $B = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$, $k = 6$

6. Sea A una matriz de orden $m \times n$ y sean I_m e I_n las matrices identidad de órdenes m y n, respectivamente.
Compruebe que

a) $A I_n = A$
b) $I_m A = A$

7. Sea A una matriz de orden $m \times n$. Demuestre que:

a) $A\, 0_{n\times p} = 0_{m\times p}$

b) $0_{q\times m} A = 0_{q\times n}$

8. Dé ejemplos de matrices cuadradas A, B y C de orden 3 tales que

a) $AB \neq BA$

b) $AB = AC$ pero $B \neq C$

c) $AB = 0$ pero $A \neq 0$ y $B \neq 0$

9. Sean A y B dos matrices ¿Qué puede concluir acerca de los órdenes de A y B, si los productos AB y BA están ambos definidos? Demuestre que una condición suficiente para que los productos AB y BA estén definidos es que A y B sean matrices cuadradas. ¿Es esta condición necesaria? Suponga ahora que los productos AB y BA están definidos y que $AB = BA$. ¿Qué puede concluir de los órdenes de A y B?

10. Sea A una matriz cuadrada. Se define A^n, $n \in N$ como

$$A^n = \underbrace{A\,A \ldots A}_{n \text{ veces}}$$

Si A es la matriz

$$A = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix}$$

compruebe que

$$A^n = \begin{bmatrix} 1+6n & 4n \\ -9n & 1-6n \end{bmatrix}$$

11. Dadas las matrices

$$A = \begin{bmatrix} \cos x & \operatorname{sen} x \\ -\operatorname{sen} x & \cos x \end{bmatrix} \qquad B = \begin{bmatrix} \cos y & \operatorname{sen} y \\ -\operatorname{sen} y & \cos y \end{bmatrix}$$

demuestre que

$$AB = BA = \begin{bmatrix} \cos(x+y) & \operatorname{sen}(x+y) \\ -\operatorname{sen}(x+y) & \cos(x+y) \end{bmatrix}$$

En particular se tiene que

$$A^2 = \begin{bmatrix} \cos 2x & \operatorname{sen} 2x \\ -\operatorname{sen} 2x & \cos 2x \end{bmatrix}$$

Demuestre que

$$A^n = \begin{bmatrix} \cos nx & \operatorname{sen} nx \\ -\operatorname{sen} nx & \cos nx \end{bmatrix}$$

12. Sea A una matriz cuadrada de orden n. Se dice que la matriz B de orden n *conmuta* con la matriz A si $AB = BA$.

a) ¿Cuáles matrices conmutan con la matriz identidad I_n?

b) ¿Cuáles matrices conmutan con la matriz cero 0_n?

c) Dada la matriz A, determine en cada caso la estructura de la matriz B de modo que B conmute con A.

c.1) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

c.2) $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$

c.3) $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

13. A la matriz A de orden n se le llama *matriz idempotente* si $A^2 = A$.

a) Compruebe que la matriz

$$A = \begin{bmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{bmatrix}$$

es idempotente.

b) Demuestre que si $AB = A$ y $BA = B$, entonces A y B son matrices idempotentes.

14. A la matriz A de orden n se le llama *matriz nilpotente* si existe un entero $k > 0$ tal que $A^k = 0$.

a) Compruebe que la matriz

$$A = \begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}$$

es nilpotente.

b) Suponga que la matriz $A = (a_{ij})_{i,j=1,\ldots,n}$ es tal que $a_{ij} = 0$ para $i \geq j$. Demuestre que A es nilpotente.

15. Se dice que la matriz A de orden n es una *matriz diagonal* si todos sus elementos, excepto los de su diagonal principal, son iguales a cero. En tal caso se escribe

$$A = \operatorname{diag}(a_1, a_2, \ldots, a_n)$$

para denotar a la matriz, diagonal

$$A = \begin{bmatrix} a_1 & & & \\ & a_2 & & 0 \\ & 0 & \ddots & \\ & & & a_n \end{bmatrix}$$

(en donde los ceros indican que todos los elementos en esa parte de la matriz son cero).

Demuestre que el producto de dos matrices (del mismo orden) diagonales es una matriz diagonal. Más concretamente, compruebe que si

$$A = \text{diag}(a_1, a_2, \ldots, a_n)$$
$$B = \text{diag}(b_1, b_2, \ldots, b_n)$$

entonces,

$$AB = BA = \text{diag}(a_1b_1, a_2b_2, \ldots, a_nb_n)$$

16. Una condición necesaria y suficiente para que la matriz B de orden n conmute con una matriz diagonal A, es que B sea una matriz diagonal. ¿Cómo tiene que ser la matriz diagonal A para que conmute con cualquier matriz B del mismo orden que A?
17. Explique por qué no es válida, en general, la siguiente fórmula, en donde A y B son matrices cuadradas del mismo orden:

$$(A + B)^2 = A^2 + 2AB + B^2$$

Dé un contraejemplo.

18. A continuación se presenta un argumento para "demostrar" que siendo A y B matrices cuadradas del mismo orden se tiene $AB = BA$.

$$AB = \frac{1}{4}(2A)(2B) = \frac{1}{4}[(A + B) + (A - B)][(A + B) - (A - B)]$$
$$= \frac{1}{4}[(A + B)^2 - (A - B)^2] = \frac{1}{4}[(B + A) - (B - A)^2]$$
$$= \frac{1}{4}[(B + A) + (B - A)][(B + A) - (B - A)] = \frac{1}{4}(2B)(2A)] = BA$$

Señale en dónde se encuentra el error de este argumento.

19. Sea A una matriz de orden $n \times p$ y B una matriz de orden $p \times m$.
    a) Demuestre que la i-ésima línea de la matriz AB se puede obtener multiplicando la i-ésima línea de la matriz A (considerada como una matriz $1 \times p$) por la matriz B.
    b) Pruebe que la j-ésima columna de la matriz AB se puede obtener multiplicando la matriz A por la j-ésima columna de la matriz B (considerada como una matriz $p \times 1$).
    c) Sean A y B las matrices

$$A = \begin{bmatrix} 1 & 4 & 3 & -1 \\ 2 & 1 & 3 & 2 \\ 1 & 1 & 4 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 0 \\ 1 & 2 & 2 \end{bmatrix}$$

Sin hacer el cálculo completo del producto AB, determine:

c.1) la primera línea de AB.

c.2) la segunda columna de AB.

c.3) la tercera columna de AB.

20. Sea A la matriz cuadrada

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

Se define *la traza* de la matriz A, denotada por trA, como

$$trA = a_{11}a_{22} + \ldots a_{nn} = \sum_{i=1}^{n} a_{ii}$$

Es decir, la traza de la matriz (cuadrada) A es la suma de los elementos de su diagonal principal.

Sean A y B las matrices

$$A = \begin{bmatrix} 2 & 5 & 4 \\ 3 & 4 & 7 \\ 0 & 1 & 9 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 7 & 4 \\ 0 & 2 & 3 \end{bmatrix}$$

a) Calcule trA, trB, $trA + trB$.

b) Calcule $tr(A + B)$ y compruebe que $tr(A + B) = trA + trB$.

c) Calcule $tr(AB)$ y $tr(BA)$. Compruebe que a pesar de que $AB \neq BA$, se tiene $tr(AB) = tr(BA)$

21. Sean A y B las matrices de orden n

$$A = (a_{ij})_{i,j=1,\ldots,n}$$
$$B = (b_{ij})_{i,j=1,\ldots,n}$$

a) Demuestre que $tr(A + B) = trA + trB$.

b) Demuestre que $tr(AB) = tr(BA)$.

c) Demuestre que si $A = B$, entonces $trA = trB$. ¿Es cierta la afirmación recíproca?

① 22. Demuestre que no existen las matrices cuadradas A y B tales que

$$AB - BA = I$$

(*Sugerencia*: proceda por contradicción. Use los resultados del ejercicio anterior.)

23. Sean A y B dos matrices del mismo orden $m \times n$ en las que se ha efectuado la misma partición. Escriba

$$A = (A_{ij})_{\substack{i=1,\ldots,m,\\ j=1,\ldots,n,}} \qquad B = (B_{ij})_{\substack{i=1,\ldots,m,\\ j=1,\ldots,n,}}$$

en donde A_{ij} y Bij son las submatrices resultantes de la partición efectuada en A y B.

a) Demuestre que la suma de las matrices A y B puede efectuarse "submatriz por submatriz". Esto es

$$A + B = (A_{ij} + B_{ij})_{\substack{i=1,\ldots,m,\\ j=1,\ldots,n,}}$$

b) Pruebe que si k es un escalar, entonces

$$kA = (kA_{ij})_{\substack{i=1,\ldots,m,\\ j=1,\ldots,n,}}$$

24. Considere las matrices cuadradas A_i y B_i de orden n_i, $i = 1, 2$. Sean A y B las matrices (de orden $n_1 + n_2$) siguientes:

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \qquad B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}$$

Demuestre que

$$A = \begin{bmatrix} A_1B_1 & 0 \\ 0 & A_2B_2 \end{bmatrix}$$

Compare con el ejercicio 15. ¿Es en este caso $AB = BA$? Calcule el producto AB si

$$A = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{bmatrix}. \qquad B = \begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 \\ 4 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 4 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 & 4 & 5 \end{bmatrix}$$

25. Sea I la matriz identidad de orden 2 y sea A la matriz

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$$

Considere las matrices e, i, j, k de orden 4

$$e = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \quad i = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} \quad j = \begin{bmatrix} A & 0 \\ 0 & -A \end{bmatrix} \quad k = \begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}$$

Demuestre que

$$ei = ie = i \qquad ij = -ji = k \qquad i^2 = -e$$
$$ej = je = j \qquad jk = -kj = i \qquad j^2 = -e$$
$$ek = ke = k \qquad ki = -ik = j \qquad k^2 = -e$$

(a una matriz de la forma $c_1e + c_2i + c_3j + c_4k$, en donde c_1, c_2, c_3 y c_4 son números reales, se le llama *cuaternio*).

26. Se dice que la matriz cuadrada A es una *matriz involutoria* si $A^2 = I$. Sea B una matriz cuadrada de orden n cualquiera. Demuestre que la matriz A de orden $2n$

$$A = \begin{bmatrix} I_n & 0 \\ B & -I_n \end{bmatrix}$$

es involutoria. Demuestre, también, la siguiente relación entre matrices involutorias y matrices idempotentes (véase ejercicio 13): Si A es involutoria, la matriz

$$C = \frac{1}{2}(A + I)$$

es idempotente. Si C es idempotente, la matriz

$$A = 2C - I$$

es involutoria.

27. En los incisos siguientes se dan matrices A y B en las que se ha efectuado cierta partición. Calcule en cada caso el producto AB atendiendo a la partición dada. Verifique su respuesta.

a) $$A = \left[\begin{array}{cc|c} 3 & 1 & 0 \\ 2 & 1 & 0 \\ \hline 0 & 0 & 4 \end{array}\right] \qquad B = \left[\begin{array}{cc|c} 5 & 8 & 24 \\ 3 & 1 & 2 \\ \hline 1 & 1 & -1 \end{array}\right]$$

b) $$A = \left[\begin{array}{ccc|c} 3 & 1 & 4 & 7 \\ 2 & 1 & 3 & 6 \\ 3 & 4 & 8 & 5 \\ \hline 0 & 7 & 1 & 2 \end{array}\right] \qquad B = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ \hline 3 & 1 & 2 & 5 \end{array}\right]$$

c) $$A = \left[\begin{array}{cc|cc} 3 & 8 & 1 & 0 \\ 4 & 5 & 0 & 1 \\ \hline 1 & 0 & 1 & -1 \\ 0 & 1 & 3 & 2 \end{array}\right] \qquad B = \left[\begin{array}{cc|cc} 5 & 2 & 4 & 5 \\ -1 & 1 & 2 & 3 \\ \hline 3 & 7 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{array}\right]$$

① 28. Sean A y B las siguientes matrices:

$$A = \begin{bmatrix} 2 & 1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & -4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 8 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 10 & 24 & 0 & -1 \end{bmatrix}$$

$$B = \begin{bmatrix} 3 & 1 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 4 & 5 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 3 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 5 & 7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 2 & 24 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 2 & -1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & 2 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & -2 & -1 & 4 & 7 & -8 & 4 & 8 & -1 & 0 \\ 3 & 2 & 2 & 7 & 7 & 4 & 10 & 24 & 0 & -1 \end{bmatrix}$$

Calcule el producto AB.

(*Sugerencia*: Una adecuada partición en las matrices A y B puede simplificar mucho los cálculos.)

29. En los incisos siguientes, se dan sistemas no homogéneos de ecuaciones lineales, los cuales son consistentes. Escriba la solución del sistema como la suma de la solución general del sistema homogéneo asociado y una solución particular del sistema no homogéneo.

a) $x_1 + x_2 = 1$
$2x_1 + 2x_2 = 2$

b) $x_1 + x_2 - x_3 = 2$
$x_1 - x_2 + x_3 = 1$

c) $x_1 + 2x_2 + x_3 = 4$
$x_1 - 2x_2 + x_3 = 0$
$2x_1 + 3x_2 - 3x_3 = 2$

d) $x_1 - x_2 + x_3 = 1$
$5x_1 + 5x_2 - x_3 = 9$
$2x_1 + 3x_2 - x_3 = 4$

30. (En un tono menos serio: un problema práctico de multiplicación de matrices.)

En el bar "La Rana Feliz" se preparan 5 tipos de bebidas cuyas recetas se muestran a continuación:

Bebida 1	(cuba libre)	1 medida de ron 1 refresco de cola
Bebida 2	(cuba campechana)	1 medida de ron $\frac{1}{2}$ refresco de cola $\frac{1}{2}$ agua mineral
Bebida 3	(cuba doble)	2 medidas de ron $\frac{1}{2}$ refresco de cola $\frac{1}{2}$ agua mineral
Bebida 4	(shandi)	$\frac{1}{2}$ cerveza $\frac{1}{2}$ limonada
Bebida 5	(especialidad de la casa)	1 medida de ron $\frac{1}{2}$ cerveza $\frac{1}{2}$ limonada

a) Estructure una matriz cuyas líneas sean "líneas de ingredientes en las bebidas" y cuyas columnas sean "columnas de tipos de bebidas".

b) La experiencia demuestra que de lunes a jueves se consumen, en promedio por día, 20 bebidas del tipo 1, 30 del tipo 2, 10 del tipo 3, 20 del tipo 4 y 30 del tipo 5, mientras que tanto el viernes como el sábado se consumen 30 bebidas del tipo 1, 40 del tipo 2, 20 del tipo 3, 40 del tipo 4 y 40 del tipo 5. Estructure una matriz cuyas líneas sean "líneas de tipos de bebidas" y cuyas columnas sean "columnas del día de la semana".

c) Use las matrices de los incisos anteriores para obtener una matriz en donde se pueda leer la cantidad de ingredientes consumidos por día. Conteste entonces las siguientes preguntas: ¿Cuántos refrescos de cola se consumen por semana?, ¿cuántas limonadas?, ¿cuántas botellas de ron, sabiendo que cada botella rinde 20 medidas?

4. MATRICES (II): INVERSIBILIDAD

En toda esta sección se van a considerar solamente las *matrices cuadradas*.

Se ha visto en la sección anterior, al hablar de multiplicación de matrices, que no es posible deducir del hecho $A \neq 0$, $AB = AC$ que $B = C$, es decir, que la ley de la cancelación no es válida para matrices.

Véase con detenimiento lo que establece esta ley en el caso de números reales. La ley dice: *si a es un número real distinto de cero y $ab = ac$, entonces $b = c$.* Recuérdese que los números reales distintos de cero son precisamente aquellos que poseen un inverso multiplicativo. Es decir, $a \neq 0$ si, y sólo si existe a^{-1} tal que $a^{-1}a = aa^{-1} = 1$. Uno se pregunta si para matrices también es cierto que A es una matriz distinta de cero si, y sólo si, existe un inverso multiplicativo para A, es decir, una matriz A^{-1} (necesariamente del mismo orden que A), tal que $A^{-1}A = AA^{-1} = I$ (la matriz identidad juega el papel del 1, como ya se había visto). Véase por ejemplo que la matriz

$$A = \begin{bmatrix} 0 & 3 \\ 0 & 2 \end{bmatrix}$$

es distinta, de cero, pero no existe ninguna matriz A^{-1} tal que $A^{-1}A = AA^{-1} = I$. Esto se puede constatar haciendo el producto de una matriz arbitraria B por A

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} 0 & 3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 3b_{11} + 2b_{12} \\ 0 & 3b_{21} + 2b_{22} \end{bmatrix} \neq I$$

(pues la matriz identidad no tiene ninguna columna de ceros).

Entonces, la propiedad de existencia de inverso multiplicativo de una matriz no es equivalente a que tal matriz sea distinta de cero (como en el caso de los números reales). Ésta es una propiedad más restrictiva que se va a estudiar en esta sección. Se llama *inversibilidad*. Obsérvese que la ley de la cancelación para matrices escrita como: *si A es una matriz para la que existe un inverso multiplicativo (en el sentido mencionado anteriormente) y si $AB = AC$, entonces $B = C$,* es válida. La demostración es: tómese la matriz A^{-1} tal que $A^{-1}A = AA^{-1} = I$. Multiplíquense ambos miembros de $AB = AC$ por A^{-1} para obtener $A^{-1}(AB) = A^{-1}(AC)$. Al usar la asociatividad del producto, se obtiene $(A^{-1}A)B = (A^{-1}A)C$, o sea, $IB = IC$, donde I es la matriz identidad y entonces finalmente, $B = C$.

4.1. DEFINICIONES Y RESULTADOS PRELIMINARES

DEFINICIÓN 4.1 Se dice que la matriz A es *inversible* si existe una matriz B (necesariamente del mismo orden que A) llamada *inversa de A*, tal que

$$BA = AB = I$$

Caso contrario se dice que A es *no inversible*.

El ejemplo trivial de matriz inversible es la matriz identidad. Otro ejemplo es la matriz

$$A = \begin{bmatrix} 1 & -4 \\ -2 & 9 \end{bmatrix}$$

A es inversible pues la matriz

$$B = \begin{bmatrix} 9 & 4 \\ 2 & 1 \end{bmatrix}$$

es tal que

$$AB = \begin{bmatrix} 1 & -4 \\ -2 & 9 \end{bmatrix}\begin{bmatrix} 9 & 4 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 4 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -4 \\ -2 & 9 \end{bmatrix} = BA$$

Una primera observación que se puede hacer, en base a la definición de inversibilidad, es que en el caso de que una matriz sea inversible, su inversa es *única*. En efecto, supóngase que B_1 y B_2 son dos inversas de A. Entonces

$$B_1 = IB_1 = (B_2A)B_1 = B_2(AB_1) = B_2I = B_2$$

↑ pues B_2 es inversa de A ↑ pues B_1 es inversa de A

A partir de este momento se puede, entonces, hablar de la inversa de la matriz (inversible) A, la cual se va a denotar por A^{-1}.

COROLARIO

Si A es inversible, su inversa es también inversible y $(A^{-1})^{-1} = A$.

DEMOSTRACIÓN De la definición de inversibilidad y la unicidad de la inversa se sigue de inmediato.

Q.E.D.

TEOREMA 4.1

Si A y B son dos matrices inversibles del mismo orden, entonces AB también es inversible y

$$(AB)^{-1} = B^{-1}A^{-1}$$

DEMOSTRACIÓN Se tiene

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

Entonces $B^{-1}A^{-1}$ es la inversa de AB.

Q.E.D.

Más generalmente, si A_i, $i = 1, \ldots, n$ son n matrices inversibles del mismo orden, entonces $A_1A_2 \ldots A_n$ es también inversible y

$$(A_1A_2 \ldots A_n)^{-1} = A_n^{-1} \ldots A_2^{-1}A_1^{-1}$$

La demostración es por inducción y se deja como ejercicio para el lector.

Se hará referencia al resultado anterior como "producto de matrices inversibles es inversible".

Sea n un número natural y A una matriz cualquiera. Se define la potencia n-ésima de la matriz A como $A^n = AA \ldots A$ (n veces A multiplicada por sí misma). Se hace el convenio de que $A^0 = I$. Si A es una matriz inversible, se puede también definir "potencias negativas"* de A, escribiendo $A^{-n} = (A^{-1})^n$. Al usar esta convención se obtiene el siguiente corolario que más que útil, ofrece una fórmula muy bonita:

COROLARIO

Si A es una matriz inversible, A^n es inversible y

$$(A^n)^{-1} = A^{-n}$$

La demostración es inmediata (véase el teorema 4.1) y se deja como ejercicio.

TEOREMA 4.2

Si A es una matriz inversible y k es un escalar distinto de cero, entonces kA también es inversible y

$$(kA)^{-1} = k^{-1}A^{-1}$$

(donde k^{-1} denota al inverso multiplicativo de k).

DEMOSTRACIÓN Se hará uso del teorema 3.2 (4).

$$(k^{-1}A^{-1})(kA) = (k^{-1}k)A^{-1}A = 1 \cdot I = I$$

$$(kA)(k^{-1}A^{-1}) = (kk^{-1})AA^{-1} = 1 \cdot I = I$$

Q.E.D.

EJEMPLO 1 Así, por ejemplo, se había visto al principio de esta subsección que la inversa de la matriz

$$A = \begin{bmatrix} 1 & -4 \\ -2 & 9 \end{bmatrix}$$

era

$$A^{-1} = \begin{bmatrix} 9 & 4 \\ 2 & 1 \end{bmatrix}$$

*Se debe tener mucho cuidado en no confundir el significado de "potencia de una matriz A" con la manera como se está denotando su inversa A^{-1}. Por ejemplo, mientras A^2 significa $A \cdot A$, A^{-1} *no significa* $1/A$, sino que es la manera como *se denota* la inversa de A.

Entonces, la inversa de la matriz

$$3A = \begin{bmatrix} 3 & -12 \\ -6 & 27 \end{bmatrix}$$

será, según el teorema 4.2, la matriz

$$\frac{1}{3}A^{-1} = \begin{bmatrix} 3 & 4/3 \\ 2/3 & 1/3 \end{bmatrix}$$

como se puede comprobar directamente.

Un resultado interesante, que conecta la propiedad de inversibilidad de la matriz A, con las soluciones del sistema homogéneo $AX = 0$, está contenido en el siguiente teorema:

TEOREMA 4.3

> Si la matriz A es inversible, entonces el sistema homogéneo de ecuaciones lineales $AX = 0$ tiene sólo la solución trivial.

DEMOSTRACIÓN Supóngase que A es inversible y considérese el sistema homogéneo $AX = 0$. Sea $\tilde{X}$ cualquier solución de este sistema, esto es, $A\tilde{X} = 0$. Multiplicando (por la izquierda) esta última igualdad por A^{-1} se obtiene $A^{-1}(A\tilde{X}) = A^{-1}0$, o sea $(A^{-1}A)\tilde{X} = 0$, esto es, $\tilde{X} = 0$. Es decir, la única solución del sistema $AX = 0$ es la trivial.

Q.E.D.

En realidad, también es cierta la afirmación recíproca del teorema anterior, es decir, si el sistema homogéneo de ecuaciones $AX = 0$ tiene sólo la solución trivial, entonces se puede concluir que la matriz A es inversible (y en ese caso decir que A es inversible es *equivalente* a decir que el sistema homogéneo de ecuaciones $AX = 0$ tiene sólo la solución trivial).

Para poder dar una demostración de este hecho se tendrá que avanzar un poco más en la teoría que se está desarrollando (la prueba se dará en la siguiente subsección). Sin embargo, se quiere usarlo ahora para dar una caracterización más simple de inversibilidad de una matriz.

TEOREMA 4.4

> Sea A una matriz cualquiera.
>
> 1) Si existe B tal que $BA = I$, entonces A es inversible (y $B = A^{-1}$).
> 2) Si existe C tal que $AC = I$, entonces A es inversible (y $C = A^{-1}$).

DEMOSTRACIÓN 1) Considérese el sistema de ecuaciones $AX = 0$. Sea $\tilde{X}$ una solución de él, esto es, $A\tilde{X} = 0$. Multiplique por la izquierda por B para obtener $\tilde{X} = 0$. Entonces, el sistema $AX = 0$ tiene sólo la solución trivial. En ese caso A es inversible. Multiplique por la derecha a $BA = I$ por A^{-1} para obtener $B = A^{-1}$.

2) Análoga a 1).

Q.E.D.

De aquí en adelante, entonces, para verificar que B es inversa de A basta con verificar que $BA = I$ o que $AB = I$ (cualquiera de estas condiciones implica a la otra, según el teorema anterior).

Hasta este momento nada se ha dicho acerca de *cómo* determinar la inversa de una matriz. De eso nos ocuparemos en la subsección 4.3. Sin embargo, se puede plantear un procedimiento natural para hallar la inversa de una matriz como se demuestra a continuación.

Sea la matriz

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Se quiere determinar bajo qué condiciones esta matriz es inversible y, en tal caso, obtener una fórmula para su inversa.

Se quiere entonces hallar una matriz

$$B = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$$

tal que

$$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Al hacer la multiplicación indicada e igualando los elementos correspondientes, se obtiene el par de sistemas de dos ecuaciones con 2 incógnitas

$$\begin{cases} ax_1 + bx_3 = 1 \\ cx_1 + dx_3 = 0 \end{cases}$$

$$\begin{cases} ax_2 + bx_4 = 0 \\ cx_2 + dx_4 = 1 \end{cases}$$

Recuérdese que a, b, c y d son dados y se quiere determinar x_1, x_2, x_3, x_4. Es fácil ver que el primer sistema tiene solución si $ad - bc \neq 0$, en cuyo caso

$$x_1 = \frac{d}{ad - bc} \qquad x_3 = \frac{-c}{ad - bc}$$

Similarmente, si $ad - bc \neq 0$, el segundo sistema tiene por solución

$$x_2 = \frac{-b}{ad - bc} \qquad x_4 = \frac{a}{ad - bc}$$

De tal manera, entonces, si $ad - bc \neq 0$, la matriz A es invertible y su inversa es:

$$A^{-1} = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} = \begin{bmatrix} \dfrac{d}{ad - bc} & \dfrac{-b}{ad - bc} \\ \dfrac{-c}{ad - bc} & \dfrac{a}{ad - bc} \end{bmatrix} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

EJEMPLO 2 Aplíquese de nuevo este procedimiento para hallar la inversa de la matriz

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}$$

Se quieren en ese caso hallar $x_1, x_2, \ldots, x_9$ tales que

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Al hacer las operaciones indicadas se obtiene

$$\begin{aligned} x_1 + 2x_4 + 3x_7 &= 1 \\ 2x_4 + 3x_7 &= 0 \\ 3x_7 &= 0 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} x_1 &= 1 \\ x_4 &= 0 \\ x_7 &= 0 \end{aligned}$$

$$\begin{aligned} x_2 + 2x_5 + 3x_8 &= 0 \\ 2x_5 + 3x_8 &= 1 \\ 3x_8 &= 0 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} x_2 &= -1 \\ x_2 &= 1/2 \\ x_8 &= 0 \end{aligned}$$

$$\begin{aligned} x_3 + 2x_6 + 3x_9 &= 0 \\ 2x_6 + 3x_9 &= 0 \\ 3x_9 &= 1 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} x_3 &= 0 \\ x_6 &= -1/2 \\ x_9 &= 1/3 \end{aligned}$$

y entonces

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$

Para una matriz de orden n, este procedimiento para hallar la inversa de la matriz conduciría a resolver n sistemas de ecuaciones con n incógnitas. Es claro que, a menos que los sistemas resultantes sean fáciles de resolver (como en el ejemplo anterior), se precisa de mucho trabajo para obtener la inversa de una matriz con este procedimiento. En la subsección 4.3 se verá un método más sistemático que requiere de menos esfuerzo. Pero antes, se tendrá que avanzar más en la teoría.

4.2. MATRICES ELEMENTALES

DEFINICIÓN 4.2 Se dice que la matriz E de orden n es una *matriz elemental* si ella puede obtenerse, a partir de la matriz identidad I_n realizando en ésta *una* operación elemental.

EJEMPLO 3 Por ejemplo, la matriz

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

es una matriz elemental, pues se obtuvo partiendo de I_3, sustituyendo en ésta su primera línea por ella misma más tres veces su tercera línea.

EJEMPLO 4 Similarmente, las matrices

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 8 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

son elementales (la primera provino de I_2 intercambiando sus líneas, la segunda también provino de I_2 multiplicando por 8 su segunda línea y la tercera provino de I_3 intercambiando su primera y tercera líneas).

Sin embargo, la matriz

$$\begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$$

no es una matriz elemental, pues para llegar a ella partiendo de I_2 hay que realizar *dos* operaciones elementales (intercambiar líneas y multiplicar la primera línea por 2).

Considere la matriz A de orden $3x3$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Sea E_1 = matriz elemental que provino de I_3 *intercambiando su primera y tercera líneas.* Realice el producto E_1A

$$E_1A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{bmatrix}$$

Observe que el resultado es la matriz que proviene de A *intercambiando su primera y tercera líneas.*

Sea E_2 = matriz elemental que provino de I_3, *sustituyendo su segunda línea por ella más k veces su tercera línea.* Realice el producto E_2A.

$$E_2A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21}+ka_{31} & a_{22}+ka_{32} & a_{23}+ka_{33} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Observe que el resultado es la matriz que proviene de A *sustituyendo su segunda línea por ella, más k veces su tercera línea.*

Sea E_3 = matriz elemental que provino de I_3 *multiplicando por $c \neq 0$ su tercera línea.* Realice el producto E_3A.

$$E_3A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ ca_{31} & ca_{32} & ca_{33} \end{bmatrix}$$

Observe que el resultado es la matriz que proviene de A *multiplicando por $c \neq 0$ su tercera línea.*

Como conclusión de estos ejemplos se ve que al multiplicar por la izquierda (se dirá "premultiplicar") la matriz A por la matriz elemental E_i, se obtiene la matriz que resulta de A al realizar en ésta la misma operación elemental que se realizó en I_3 para obtener E_i (i=1, 2, 3).

Éste es en realidad un hecho general que se enuncia en el siguiente teorema, cuya demostración se deja como ejercicio al lector:

TEOREMA 4.5

> Sea A una matriz $m \times n$. Al premultiplicar A por la matriz elemental E se obtiene por resultado la matriz que proviene de A realizando en ésta la misma operación elemental efectuada en I_m para llegar a E.

Antes de enunciar y probar el siguiente resultado, que indicará que todas las matrices elementales son inversibles, se recuerda que las operaciones elementales en las líneas de una matriz tienen una operación elemental asociada "inversa" (en el sentido de que deshace la operación efectuada), que además es del mismo tipo.

En efecto,

1) Supóngase que A' se obtuvo de A intercambiando las líneas i y j. Si en A' se intercambian nuevamente las líneas i y j se recupera A.

2) Supóngase que A' se obtuvo de A sustituyendo la i-ésima línea por ella misma más k veces la j-ésima línea. Si en A' se sustituye la i-ésima línea más $(-k)$ veces la j-ésima línea, se recupera A.

Por último:

3) Si A' se obtuvo de A multiplicando su i-ésima línea por la constante $c \neq 0$, multiplíquese la i-ésima línea de A' por $\frac{1}{c}$ y recupérese así A.

TEOREMA 4.6

> Si E es una matriz elemental, entonces E es inversible. Su inversa es también una matriz elemental.

DEMOSTRACIÓN Siendo E una matriz elemental, E provino de I realizando en ésta una operación elemental. Según la observación previa al teorema, se puede recuperar I realizando en E la operación elemental inversa. Sea E' la matriz elemental que proviene de I realizando en ésta la operación elemental inversa mencionada. Según el teorema 4.5 se tiene $E'E = I$. Un argumento similar aplicado a la matriz elemental E', muestra que existe una matriz elemental E'' tal que $E''E' = I$. De las expresiones $E'E = I = E''E'$ se concluye que $E = E''$. Se ha entonces probado que $E'E = EE' = I$, esto es, que E es inversible y que su inversa E' es también una matriz elemental.

Q.E.D.

EJEMPLO 5 Por ejemplo, si E es la matriz elemental que proviene de I_3 sustituyendo su primera línea por ella misma más 3 veces su tercera línea, o sea

$$E = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

su inversa E^{-1} será aquella matriz elemental que proviene de I_3 realizando la operación elemental inversa a la que se realizó en I_3 para obtener E; es decir, sustituyendo su primera línea por ella misma más (-3) veces su tercera línea, o sea

$$E^{-1} = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

El siguiente teorema (4.7) es un resultado muy importante, tanto desde el punto de vista teórico, como también desde el punto de vista práctico, pues de él (de su demostración) se podrá deducir un *método* para hallar la inversa de una matriz. Para demostrarlo, se pondrán en juego todas las ideas desarrolladas en esta

subsección, así como también el siguiente resultado "técnico" sobre sistemas de ecuaciones.

LEMA

> Sea A una matriz de orden n. Si el sistema homogéneo de ecuaciones $AX = 0$ tiene sólo la solución trivial, entonces A es equivalente a la matriz identidad I_n.

DEMOSTRACIÓN Basta observar que el sistema $AX = 0$ es equivalente al sistema $x_1 = 0$, $x_2 = 0, \ldots, x_n = 0$, cuya matriz (la matriz del sistema) es I_n. Lo cual significa que partiendo de A, por medio de operaciones elementales, se puede llegar a I_n. Esto es precisamente lo que se quería probar.

Q.E.D.

TEOREMA 4.7

> Las siguientes afirmaciones acerca de la matriz A de orden n son equivalentes:
>
> 1) A es inversible.
> 2) A es equivalente a la matriz identidad I_n.

DEMOSTRACIÓN 1) $\Rightarrow$ 2): Si A es inversible, por el teorema 4.3, el sistema homogéneo de ecuaciones $AX = 0$ tiene sólo la solución trivial. Entonces, según el lema anterior, A es equivalente a I_n.

2) $\Rightarrow$ 1): Que A es equivalente a la matriz identidad I_n, significa que, partiendo de A y realizando en ella operaciones elementales, se puede llegar (dígase que después de r pasos) a I. Esquemáticamente

$$A \overset{e_1}{\sim} A_1 \overset{e_2}{\sim} A_2 \sim \ldots \overset{e_r}{\sim} I$$

donde e_i, $i = 1, \ldots, r$ indica que se efectúa una operación elemental. Por el teorema 4.5 se tiene que existen matrices elementales $E_1, E_2, \ldots, E_r$ tales que $E_1A = A_1$, $E_2A_1 = A_2, \ldots, E_rA_{r-1} = I$. Combinando estas expresiones se llega a $(E_r \ldots E_2E_1)A = I$. Siendo $E_1, E_2, \ldots, E_r$ inversibles (teorema 4.6) se puede presentar la expresión anterior como $A = E_1^{-1} E_2^{-1} \ldots E_r^{-1}$. Entonces A es un producto de matrices inversibles. Es por tanto inversible.

Q.E.D.

Se puede ahora probar el resultado que se había prometido en la subsección anterior.

COROLARIO

> Sea A una matriz de orden n. Si el sistema de ecuaciones $AX = 0$ tiene sólo la solución trivial, entonces A es inversible.

DEMOSTRACIÓN Según el lema previo al teorema 4.7, A es equivalente a la matriz identidad I_n. Entonces, según el teorema 4.7, A es inversible.

Q.E.D.

4.3. EVALUACIÓN DE A^{-1}

En la subsección anterior se desarrolló más la teoría acerca de la inversibilidad de matrices. Se obtuvieron resultados importantes (como el teorema 4.7) que se usarán posteriormente en otros capítulos. Sin embargo, se verá que, como fruto de esta maquinaria teórica desarrollada, se puede *concretar* un método para evaluar la inversa de una matriz.

Se centra la atención en la demostración de la implicación 2) $\Rightarrow$ 1) en el teorema 4.7. Bajo la hipótesis de que A es equivalente a la matriz identidad I_n, se concluyó que A es inversible.

Al premultiplicar la matriz A por la matriz elemental E_1 se obtiene A_1, que es la primera matriz equivalente a A que aparece en nuestro camino para llegar a I_n. Similarmente al premultiplicar A_1 por E_2 se obtiene A_2, segunda matriz, equivalente a A en el proceso para alcanzar I_n, etc. Se tiene la importante fórmula

$$(E_rE_{r-1} \ldots E_2E_1)A = I \qquad \text{(a)}$$

de la cual se puede concluir que A es inversible.

Acontece que en realidad *se puede concluir más* todavía. Tal fórmula nos está diciendo *cuál es la matriz* A^{-1}. En efecto, según el teorema 4.4 se tiene

$$A^{-1} = E_rE_{r-1} \ldots E_2E_1I \qquad \text{(b)}$$

De esta fórmula, por simple que parezca, se puede deducir un ingenioso e interesante método para hallar la inversa de una matriz.

Se ha de leer la fórmula (a) de la siguiente manera: "al efectuar las operaciones elementales correspondientes a $E_1, E_2, \ldots, E_r$ (en este orden) en la matriz A, se obtiene I".

Al leer entonces de la misma manera la fórmula (b) se tiene: "al efectuar las operaciones elementales correspondientes a $E_1, E_2 \ldots, E_r$ en la matriz I, se obtiene A^{-1}".

¡. . . y he aquí el método para hallar A^{-1}!: las mismas operaciones elementales que se hagan en A para obtener I, se hacen en I y el resultado es ni más ni menos que A^{-1}.

En la práctica, para hallar A^{-1} se va a escribir una matriz de la forma

$$[A \mid I]$$

Al centrar la atención en la parte izquierda de esta matriz (en donde se encuentra A) se realizan las operaciones elementales que conduzcan a llevar a A a la matriz identidad (obsérvese entonces que esas mismas operaciones se están efectuando en I). Cuando se llegue a I, la matriz que aparezca en el lado derecho debe ser A^{-1}.

Esquemáticamente:

$$[A \mid I] \quad \ldots \quad [I \mid A^{-1}]$$

Considérese que si llegara a aparecer una línea de ceros en la parte izquierda de la matriz durante el proceso descrito, esto indicará que A no es equivalente a I_n, y por lo tanto, según el teorema 4.7, que en tal caso A no es inversible.

Es decir, este método ayuda a detectar también la no inversibilidad de una matriz.

EJEMPLO 6 Observe funcionar este método en un ejemplo concreto. Sea

$$A = \begin{bmatrix} 3 & 2 & -1 \\ -1 & 2 & 3 \\ -3 & 1 & 3 \end{bmatrix}$$

Se irá escribiendo en la parte izquierda la matriz obtenida y en la parte derecha las operaciones elementales que se realizaron para obtenerla (de la anterior)

$$\left[\begin{array}{ccc|ccc} 3 & 2 & -1 & 1 & 0 & 0 \\ -1 & 2 & 3 & 0 & 1 & 0 \\ -3 & 1 & 3 & 0 & 0 & 1 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} & 0 & 0 \\ -1 & 2 & 3 & 0 & 1 & 0 \\ -3 & 1 & 3 & 0 & 0 & 1 \end{array}\right] \quad (L1) \to \frac{1}{3}(L1)$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{8}{3} & \frac{8}{3} & \frac{1}{3} & 1 & 0 \\ 0 & 3 & 2 & 1 & 0 & 1 \end{array}\right] \quad \begin{array}{l} (L2) \to (L2) + (L1) \\ (L3) \to (L3) + 3(L1) \end{array}$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 1 & \frac{1}{8} & \frac{3}{8} & 0 \\ 0 & 3 & 2 & 1 & 0 & 1 \end{array}\right] \quad (L2) \to \frac{3}{8}(L2)$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 1 & \frac{1}{8} & \frac{3}{8} & 0 \\ 0 & 0 & -1 & \frac{5}{8} & -\frac{9}{8} & 1 \end{array}\right] \quad (L3) \to (L3) - 3(L2)$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 1 & \frac{1}{8} & \frac{3}{8} & 0 \\ 0 & 0 & 1 & -\frac{5}{8} & \frac{9}{8} & -1 \end{array}\right] \quad (L3) \to -(L3)$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & \frac{2}{3} & 0 & \frac{1}{8} & \frac{3}{8} & -\frac{1}{3} \\ 0 & 1 & 0 & \frac{3}{4} & -\frac{3}{4} & 1 \\ 0 & 0 & 1 & -\frac{5}{8} & \frac{9}{8} & -1 \end{array}\right] \quad \begin{array}{l} (L1) \to (L1) + \frac{1}{3}(L3) \\ (L2) \to (L2) - (L3) \end{array}$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{3}{8} & \frac{7}{8} & -1 \\ 0 & 1 & 0 & \frac{3}{4} & -\frac{3}{4} & 1 \\ 0 & 0 & 1 & -\frac{5}{8} & \frac{9}{8} & -1 \end{array}\right] \quad (L1) \to (L1) - \frac{2}{3}(L2)$$

Entonces

$$A^{-1} = \begin{bmatrix} -\frac{3}{8} & \frac{7}{8} & -1 \\ \frac{3}{4} & -\frac{3}{4} & 1 \\ -\frac{5}{8} & \frac{9}{8} & -1 \end{bmatrix}$$

EJEMPLO 7 Observe un ejemplo más. Considere la matriz

$$A = \begin{bmatrix} 1 & 2 & -1 & 4 \\ 3 & 2 & 1 & 2 \\ 5 & 6 & -1 & 10 \\ 3 & -2 & 5 & 1 \end{bmatrix}$$

Obtenga

$$\left[\begin{array}{cccc|cccc} 1 & 2 & -1 & 4 & 1 & 0 & 0 & 0 \\ 3 & 2 & 1 & 2 & 0 & 1 & 0 & 0 \\ 5 & 6 & -1 & 10 & 0 & 0 & 1 & 0 \\ 3 & -2 & 3 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{cccc|cccc} 1 & 2 & -1 & 4 & 1 & 0 & 0 & 0 \\ 0 & -4 & 4 & -10 & -3 & 1 & 0 & 0 \\ 0 & -4 & 4 & -10 & -5 & 0 & 1 & 0 \\ 0 & -8 & 8 & -11 & -3 & 0 & 0 & 1 \end{array}\right] \quad \begin{array}{l} (L2) \to (L2) - 3(L1) \\ (L3) \to (L3) - 5(L1) \\ (L4) \to (L4) - 3(L1) \end{array}$$

$$\sim \left[\begin{array}{cccc|cccc} 1 & 2 & -1 & 4 & 1 & 0 & 0 & 0 \\ 0 & -4 & 4 & -10 & -3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & -1 & 1 & 0 \\ 0 & -8 & 8 & -11 & -3 & 0 & 0 & 1 \end{array}\right] \quad (L3) \to (L3) - (L2)$$

Se ha obtenido una línea de ceros en la matriz de la izquierda.

La matriz A es, pues, no inversible.

4.4. MÁS SOBRE INVERSIBILIDAD

Para terminar esta sección, se va a explotar un poco más la idea del método que se acaba de describir, para obtener algunos otros resultados interesantes.

TEOREMA 4.8 La matriz A es inversible si, y sólo si, se puede escribir como producto de matrices elementales.

DEMOSTRACIÓN Si la matriz A se puede escribir como producto de matrices elementales, es claro que A es inversible, pues las matrices elementales lo son y el producto de matrices inversibles es inversible. Recíprocamente, supóngase que existe A^{-1}. La fórmula (b) dice que

$$A^{-1} = E_r E_{r-1} \dots E_2 E_1$$

donde $E_1, E_2, \dots, E_r$ son matrices elementales. Entonces

$$A = (A^{-1})^{-1} = (E_r E_{r-1} \dots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \dots E_r^{-1}$$

Las matrices E_i^{-1} $(i=1, 2, \dots, r)$ son inversas de matrices elementales, y son, por tanto (véase teorema 4.6), matrices elementales.

Q.E.D.

EJEMPLO 8 Por ejemplo, considere la matriz

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 6 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Observe que A es inversible

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 3 & 6 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -3 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \quad (L2) \to (L2) - 3(L1)$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -3 & 1 & 0 \end{array}\right] \quad (L2) \leftrightarrow (L3)$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & -2 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -3 & 1 & 0 \end{array}\right] \quad (L1) \to (L1) - 2(L2)$$

Entonces

$$A^{-1} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 0 & 1 \\ -3 & 1 & 0 \end{bmatrix}$$

Según el teorema anterior A debe de expresarse como el producto de matrices elementales. Aún más, el mismo teorema dice que $A = E_1^{-1}E_2^{-1} \dots E_r^{-1}$, donde $E_1^{-1}E_2^{-1}, \dots, E_r^{-1}$ representan las matrices elementales inversas de las matrices correspondientes a las operaciones elementales que se tuvieron que realizar en A para llegar a I. En nuestro ejemplo se tiene:

OPERACIÓN ELEMENTAL	E	E^{-1}
$(L2) \to (L2) - 3(L1)$	$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$(L2) \leftrightarrow (L3)$	$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	$E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
$(L1) \to (L1) - 2(L2)$	$E_3 = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$E_3^{-1} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Entonces

$$A = E_1^{-1}E_2^{-1}E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

o sea,

$$\begin{bmatrix} 1 & 2 & 0 \\ 3 & 6 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Otro resultado interesante, de carácter más general que el anterior, está contenido en el siguiente teorema:

TEOREMA

Sea A una matriz $m \times n$. Entonces A puede factorizarse como $A = QA'$, donde A' es la forma escalonada reducida de A y Q es una matriz inversible (de orden m).

DEMOSTRACIÓN La idea de la demostración es la misma que se usó para demostrar 2) ⟹ 1) en el teorema 4.7. Partiendo de A, por medio de operaciones elementales, se llega a A'. Cada operación elemental que se realiza en A equivale a premultiplicar por una matriz elemental. Existen pues, k matrices elementales $E_1, E_2, \ldots, E_k$ de orden m tales que $E_k \ldots E_2E_1A = A'$. Entonces $A = (E_1^{-1} \ldots E_k^{-1})A'$. Llámese $Q = E_1^{-1} \ldots E_k^{-1}$. Es claro que Q es inversible.

Q.E.D.

El mismo teorema anterior proporciona el método para hallar la factorización mencionada de una matriz. Esquemáticamente el método se ve como

$$[A_{m\times n} \mid I_m] \sim \quad \ldots \quad \sim [A' \mid Q^{-1}]$$

EJEMPLO 9

Por ejemplo, si A es la matriz

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 4 & 5 \end{bmatrix}$$

se tiene

$$\left[\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ -1 & 4 & 5 & 0 & 1 \end{array}\right] \sim \ldots \sim \left[\begin{array}{ccc|cc} 1 & 0 & 1/3 & 2/3 & -1/3 \\ 0 & 1 & 4/3 & 1/6 & 1/6 \end{array}\right]$$

Entonces,

$$A' = \begin{bmatrix} 1 & 0 & 1/3 \\ 0 & 1 & 4/3 \end{bmatrix}$$

Ahora obtenga $Q = (Q^{-1})^{-1}$

$$\left[\begin{array}{cc|cc} 2/3 & -1/3 & 1 & 0 \\ 1/6 & 1/6 & 0 & 1 \end{array}\right] \sim \ldots \sim \left[\begin{array}{cc|cc} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 4 \end{array}\right]$$

Entonces

$$Q = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$$

La factorización de A se ve entonces como

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ -1 & 4 & 5 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}}_{\text{inversible}} \underbrace{\begin{bmatrix} 1 & 0 & 1/3 \\ 0 & 1 & 4/3 \end{bmatrix}}_{\text{forma escalonada reducida de } A}$$

Un último resultado que relaciona la propiedad de inversibilidad de una matriz A con las soluciones del sistema no homogéneo de ecuaciones lineales es:

TEOREMA 4.10

Las siguientes afirmaciones sobre la matriz A de orden n son equivalentes:

1) A es inversible.
2) El sistema $AX = B$ tiene solución única.

DEMOSTRACIÓN 1) ⟹ 2). Siendo A inversible, el sistema $AX = B$ tiene al menos la solución $\tilde{X} = A^{-1}B$ (pues $A(A^{-1}B) = B$ es una identidad). Para ver que esta solución es única, sea $\tilde{X}'$ cualquier otra solución del sistema. Por lo tanto, $A\tilde{X}' = B$. Premultiplique por A^{-1} y obtenga $\tilde{X}' = A^{-1}B = \tilde{X}$ (otro argumento para la unicidad: sea $\tilde{X}$ cualquier solución del sistema $AX = B$. Entonces $\tilde{X} = \tilde{X}_p + \tilde{X}_h$ donde $\tilde{X}_p$ es *una* solución particular al sistema y $\tilde{X}_h$ es la solución general del sistema homogéneo asociado. Pero A inversible implica $\tilde{X}_h = 0$ (teorema 4.3). Entonces, $\tilde{X} = \tilde{X}$ y la solución es única).

2) ⟹ 1). La demostración del corolario del teorema 3.3 permite concluir que si $AX = B$ tiene una solución única, entonces el sistema homogéneo asociado tiene sólo la solución trivial. De este modo, por el corolario del teorema 4.7, la matriz A es inversible.

Q.E.D.

Este teorema proporciona un nuevo método para resolver sistemas no homogéneos de ecuaciones "cuadrados" (número de ecuaciones = número de incógnitas) cuando éstos tienen una solución única.

EJEMPLO 10

Por ejemplo, considere el sistema

$$\begin{aligned} 3x_1 + 2x_2 - x_3 &= 4 \\ -x_1 + 2x_2 + 3x_3 &= 1 \\ -3x_1 + x_2 + 3x_3 &= 2 \end{aligned}$$

o bien, $AX = B$ con

$$A = \begin{bmatrix} 3 & 2 & -1 \\ -1 & 2 & 3 \\ -3 & 1 & 3 \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad B = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$$

ya se vio (primer ejemplo de la subsección 4.3) que A es inversible y que

$$A^{-1} = \begin{bmatrix} -3/8 & 7/8 & -1 \\ 3/4 & -3/4 & 1 \\ -5/8 & 9/8 & -1 \end{bmatrix}$$

Por consiguiente, la solución (única) del sistema está dada por

$$\bar{X} = A^{-1}B = \begin{bmatrix} -3/8 & 7/8 & -1 \\ 3/4 & -3/4 & 1 \\ -5/8 & 9/8 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -21/8 \\ 17/4 \\ -27/8 \end{bmatrix}$$

es decir, $x_1 = -21/8$, $x_2 = 17/4$, $x_3 = -27/8$.

EJERCICIOS (SECCIÓN 4, CAPÍTULO 1)

1. Verifique que la matriz

$$B = \begin{bmatrix} 5 & -4 & -6 \\ -9 & 8 & 11 \\ 1 & -1 & -1 \end{bmatrix}$$

es la inversa de la siguiente matriz A. ¿Cuál es la inversa de B?

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 1 & -1 \\ 1 & 1 & 4 \end{bmatrix}$$

2. Demuestre que la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 10 & 20 & -1 & 0 \\ 5 & -8 & 0 & -1 \end{bmatrix}$$

es inversible y que ella es su propia inversa. (Ver ejercicio 25 de la sección anterior.)

3. Demuestre que una matriz con una línea o con una columna de ceros no puede ser inversible.

4. Verifique la validez del teorema 4.1 con las matrices A del ejercicio 1 y la matriz

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}$$ (Cuya inversa fue determinada en la página 86.)

5. Demuestre el recíproco del teorema 4.1: si A y B son dos matrices del mismo orden tales que el producto AB es inversible, entonces A y B son matrices inversibles. (Observe que *no* puede usar la fórmula $(AB)^{-1} = B^{-1}A^{-1}$, pues ésta presupone precisamente lo que se quiere demostrar.)

6. Sea A la matriz

$$A = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix}$$

Compruebe que A es inversible y halle su inversa. En el ejercicio 10 de la sección anterior se demostró que

$$A^n = \begin{bmatrix} 1+6n & 4n \\ -9n & 1-6n \end{bmatrix}$$

en donde $n = 1, 2, \ldots$ Pruebe que esta fórmula también es válida para "potencias negativas de A".

7. Encuentre la inversa de la matriz

$$B = \begin{bmatrix} 15 & -12 & -18 \\ -27 & 24 & 33 \\ 3 & -3 & -3 \end{bmatrix}$$

(*Sugerencia*: Use el teorema 4.2 y el ejercicio 1.)

8. Si A es una matriz inversible y k es un número distinto de cero, demuestre que kA^n y $(kA)^n$ son inversibles, en donde n es un número natural cualquiera.

9. Sean $A_1, A_2, \ldots, A_n$ matrices inversibles y $k_1, k_2, \ldots, k_n$ números no nulos. Demuestre (por inducción) que el producto

$$(k_1A_1)(k_2A_2)\ldots(k_nA_n)$$

es inversible y que su inversa es

$$\frac{1}{k_1k_2\ldots k_n} A_n^{-1}A_{n-1}^{-1}\ldots A_1^{-1}$$

10. Use el ejercicio 1 y el teorema 4.3 para resolver el sistema homogéneo de ecuaciones lineales

$$5x_1 - 4x_2 - 6x_3 = 0$$

$$-9x_1 + 8x_2 + 11x_3 = 0$$
$$x_1 - x_2 - x_3 = 0$$

11. En esta sección (pág. 86) se demostró que si $ad - bc \neq 0$ entonces la matriz

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

es inversible. Demuestre la afirmación recíproca: si A es inversible entonces $ad - bc \neq 0$.

12. Utilice el "método" descrito en la página 85 (para hallar inversas de matrices) para demostrar que la matriz

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 5 & 4 \end{bmatrix}$$

no es inversible.

13. Dada la matriz

$$A = \begin{bmatrix} 3 & -1 & 4 \\ 0 & 2 & 1 \\ 1 & -1 & -2 \end{bmatrix}$$

demuestre que $A^3 - 3A^2 - 7A + 18I_3 = 0$ (la matriz cero de orden 3)

A partir de este resultado pruebe que A es inversible y halle A^{-1}.

14. Dada la matriz

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 1 \\ 2 & 3 & 1 \end{bmatrix}$$

compruebe que $A^3 - 3A^2 - 7A - 11I_3 = 0$. Encuentre A^{-1}.

15. Diga cuáles de las siguientes matrices, son matrices elementales. En tal caso, encuentre su inversa usando el teorema 4.6 (su demostración).

a) $\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ d) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix}$

16. Considere las siguientes matrices elementales:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_2 = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Sea A una matriz cualquiera de orden 3. Describa las siguientes matrices:

a) E_1A	d) E_1E_3A	g) E_2E_3A	j) $E_3^2E_2A$
b) E_2A	e) E_1E_2A	h) E_3E_1A	k) $E_2^{-1}E_3^2E_1A$
c) E_3A	f) E_2E_1A	i) $E_3E_1E_2A$	l) $E_1^{-1}E_2E_3E_2^{-1}A$

17. Sea A una matriz $m \times p$ y B una matriz $p \times n$. ¿Cómo se altera el producto AB si en la matriz A
   a) intercambia dos de sus líneas,
   b) multiplica una de sus líneas por el escalar $k \neq 0$,
   c) sustituye una de sus líneas por ella misma más k veces otra línea

   (*Sugerencia*: use el teorema 4.5 y la propiedad asociativa del producto de matrices.)

18. En esta sección se discutió cuál es el efecto de premultiplicar una matriz A por una matriz elemental E (teorema 4.5). ¿Qué ocurre si a la matriz A la multiplicamos *por la derecha* por la matriz elemental E? Es decir, ¿cómo es la matriz AE respecto de la matriz A?

19. Sea A una matriz $m \times p$ y B una matriz $p \times n$. ¿Cómo se altera el producto AB si en la matriz B,
   a) intercambia dos de sus columnas,
   b) multiplica una de sus columnas por el escalar $k \neq 0$,
   c) sustituye una de sus líneas por ella misma más k veces otra columna

① 20. Sea A una matriz inversible. Describa los cambios que ocurren en A^{-1} si en la matriz A
   a) se intercambian dos de sus líneas,
   b) se multiplica una de sus líneas por el escalar $k \neq 0$,
   c) se sustituye una de sus líneas por ella misma más k veces otra línea.

   (*Sugerencia*: use los teoremas 4.6 y 4.1, así como el resultado del ejercicio anterior.)

21. Repita el ejercicio anterior cambiando la palabra "línea" por la palabra "columna".

22. Use el teorema 4.7 para demostrar que la matriz

$$A = \begin{bmatrix} 8 & 5 & 4 \\ 0 & 7 & 1 \\ 0 & 0 & -2 \end{bmatrix}$$

es inversible.

23. Sea A la matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & a_{22} & \dots & a_{2n} \\ & & \ddots & \\ 0 & & & a_{nn} \end{bmatrix}$$

en donde $a_{11}a_{22}\ldots a_{nn} \neq 0$. Demuestre que A es inversible. ¿Qué acontece si $a_{11}a_{22}\ldots a_{nn} = 0$?

(*Sugerencia*: use el mismo argumento que en el ejercicio anterior.)

24. Como un caso particular del ejercicio anterior, concluya que la matriz diagonal

$$A = \operatorname{diag}(a_1, a_2, \ldots, a_n)$$

en donde $a_1a_2\ldots a_n \neq 0$, es inversible. Calcule su inversa.

(*Sugerencia*: véase ejercicio 15 de la sección anterior.)

25. Sea A una matriz nilpotente (ver ejercicio 14 de la sección anterior). Compruebe que A no puede ser inversible.
26. Sea A una matriz involutoria (ver ejercicio 26 de la sección anterior). Demuestre que A es inversible. Determine su inversa.
27. Pruebe que si A es una matriz idempotente (ver ejercicio 13 de la sección anterior) e inversible, entonces A es la matriz identidad.
28. Sean A y B dos matrices inversibles de orden n. Demuestre que la matriz C de orden $2n$

$$C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

es inversible. Determine su inversa. Use este hecho para hallar la inversa de la matriz

$$C = \begin{bmatrix} 7 & 3 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

(*Sugerencia*: véase el ejercicio 23 de la sección anterior.)

29. Diga cuáles de las siguientes matrices son inversibles. En caso de que lo sean, halle su inversa

a) $\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 4 & 4 \\ 2 & 3 \end{bmatrix}$

c) $\begin{bmatrix} 3 & 2 \\ 6 & 4 \end{bmatrix}$

d) $\begin{bmatrix} 1 & 2 & 8 \\ 0 & 6 & 9 \\ 0 & 2 & 3 \end{bmatrix}$

e) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & -1 & 4 \end{bmatrix}$

f) $\begin{bmatrix} 1 & 0 & 2 \\ -2 & 3 & 1 \\ 0 & 3 & 5 \end{bmatrix}$

g) $\begin{bmatrix} 1 & 5 & 3 \\ 2 & 7 & 3 \\ 3 & 9 & 4 \end{bmatrix}$

h) $\begin{bmatrix} 1 & 2 & 3 & -2 \\ 0 & 1 & 2 & 3 \\ -1 & -2 & 0 & -1 \\ 2 & 1 & -1 & 1 \end{bmatrix}$

i) $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

j) $\begin{bmatrix} 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ -2 & 2 & -1 & 1 \end{bmatrix}$

k) $\begin{bmatrix} 2 & 3 & 1 & 2 \\ 1 & 1 & 1 & -1 \\ 1 & 2 & 3 & 4 \\ 1 & 0 & -2 & -6 \end{bmatrix}$

l) $\begin{bmatrix} -1 & 2 & 0 & 3 & 1 \\ 1 & 0 & 3 & -2 & 2 \\ 2 & 0 & 0 & -4 & -2 \\ 0 & 1 & 0 & 2 & -3 \\ 0 & -1 & 2 & 1 & 1 \end{bmatrix}$

m) $\begin{bmatrix} 1 & 1 & 0 & 0 & -1 & 2 \\ 1 & 0 & 2 & 1 & 3 & 5 \\ 0 & 1 & 2 & -1 & 0 & 1 \\ -2 & 0 & -1 & -2 & 0 & -6 \\ 3 & -1 & -2 & 0 & -5 & -9 \\ 0 & 0 & -1 & 2 & -2 & 3 \end{bmatrix}$

30. (El mensaje secreto.) En este problema se da un mensaje para todas aquellas personas que estudian este libro. Considere la equivalencia letras-números: $A = 1$, $B = 2$, $C = 3$, $\ldots$, $Z = 27$. Se proporciona la matriz BA de orden 4×8 siguiente:

$$BA = \begin{bmatrix} 33 & 18 & 70 & 29 & 60 & 24 & 16 & 36 \\ 165 & 94 & 130 & 117 & 112 & 106 & 70 & 184 \\ 12 & 34 & -46 & -10 & -46 & 21 & 15 & 10 \\ 68 & 48 & 76 & 30 & 65 & 50 & 30 & 52 \end{bmatrix}$$

La llave que descubre el mensaje es la matriz

$$B = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 3 & 4 & 0 & 3 \\ -1 & 2 & -1 & -1 \\ 0 & 1 & 3 & 2 \end{bmatrix}$$

En las líneas de la matriz A se lee el mensaje. Si lo descubre, póngalo en práctica.

31. En cada uno de los ejercicios siguientes, X es una matriz de orden 2. Resuelva para X.

a) $\begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} X = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$

b) $\begin{bmatrix} 4 & 1 \\ -3 & -1 \end{bmatrix} X = \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} X \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}$

d) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^3 X \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}^2 = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$

32. En cada uno de los ejercicios siguientes, X es una matriz de orden 3. Resuelva para X.

a) $$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 1 \\ 3 & 2 & 3 \end{bmatrix} X = \begin{bmatrix} 1 & 3 & 2 \\ -1 & 0 & 4 \\ 2 & 1 & 1 \end{bmatrix}$$

b) $$\begin{bmatrix} 4 & 5 & 3 \\ 0 & 1 & 0 \\ 2 & 8 & 4 \end{bmatrix}^2 X = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

c) $$\begin{bmatrix} 1 & 3 & 1 \\ 2 & 2 & 3 \\ 0 & 3 & 2 \end{bmatrix} X \begin{bmatrix} 1 & 2 & -3 \\ 2 & -1 & 0 \\ 3 & 2 & -4 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix}$$

d) $$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}^4 X \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}^3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

33. Demuestre que la matriz

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 5 \end{bmatrix}$$

es inversible y escríbala como un producto de matrices elementales.

34. En el ejercicio 24 se demostró que la matriz diagonal

$$A = \text{diag}\,(a_1, a_2, \ldots, a_n)$$

en donde $a_1, a_2 \ldots a_n \neq 0$ es inversible. Escriba esta matriz como un producto de n matrices elementales.

35. Según el teorema 4.9 toda matriz A se puede escribir como el producto QA' en donde Q es una matriz inversible y A' es la forma escalonada reducida de A. ¿Cómo se ve este producto si A es una matriz inversible?

36. Escriba cada una de las siguientes matrices como el producto de una matriz inversible Q y la forma escalonada reducida de A:

a) $\begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 4 \end{bmatrix}$ b) $\begin{bmatrix} 3 & 1 & 4 \\ 2 & 1 & 3 \\ -1 & -1 & -2 \end{bmatrix}$ c) $\begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 1 & 1 \end{bmatrix}$

37. Resuelva los siguientes sistemas de ecuaciones, invirtiendo primeramente la matriz del sistema y multiplicando luego la inversa por la matriz de términos independientes

(véase demostración del teorema 4.10):

a) $$\begin{aligned} 2x_1 + 3x_2 &= 1 \\ 3x_1 + 4x_2 &= 2 \end{aligned}$$

b) $$\begin{aligned} 3x_1 - 3x_2 + 2x_3 &= 1 \\ 4x_1 - 5x_2 + 2x_3 &= 1 \\ 5x_1 - 6x_2 + 4x_3 &= 3 \end{aligned}$$

c) $$\begin{aligned} x_1 + x_2 - x_3 + x_4 &= 3 \\ 2x_1 - 3x_2 + 4x_3 - 2x_4 &= 1 \\ 3x_1 - x_2 - 2x_3 - 2x_4 &= -2 \\ x_1 + x_2 - 2x_3 + 5x_4 &= 5 \end{aligned}$$

CAPÍTULO DOS

Determinantes

El objetivo principal de este capítulo es crear una herramienta adicional de trabajo que ayudará a avanzar en la teoría que se desarrollará en los próximos capítulos. Esta herramienta está relacionada con la teoría de determinantes. Se debe decir, sin embargo, que la teoría de determinantes tiene por sí misma un carácter propio, independientemente del uso que de ella se hace como herramienta de trabajo en el álgebra lineal. En este capítulo se pretende también ver algunos aspectos —elementales— de esta interesante teoría.

Uno de los conceptos más importantes que han aparecido hasta este momento en el libro, es el concepto de inversibilidad de una matriz. Relacionado con él, aparecerán posteriormente muchas otras ideas y resultados. En este capítulo se verá que con la ayuda de los determinantes es posible "etiquetar" a las matrices cuadradas con información sobre la inversibilidad de éstas. Más concretamente, se establecerá una función *det*: $M_{n \times n} \to R$ llamada (función) *determinante* del conjunto de matrices cuadradas de orden n a los números reales, tal que $det(A) \in R$ —valor que se calcula con los coeficientes de A— dirá si la matriz A es inversible o no lo es.

En el capítulo anterior se vio que si $ad - bc \neq 0$, la matriz

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

es inversible. Resultará que el número $ad - bc \in R$ es precisamente el determinante de la matriz A. Se verá que en general una condición necesaria y suficiente para que una matriz cuadrada A sea inversible es que su determinante sea distinto de cero (corolario 1 del teorema 3.1).

Para poder abordar adecuadamente la teoría que se desarrollará sobre determinantes, se deben desarrollar previamente algunas ideas sobre permutaciones. Éste es el objetivo de la primera sección que ahora se comienza a estudiar.

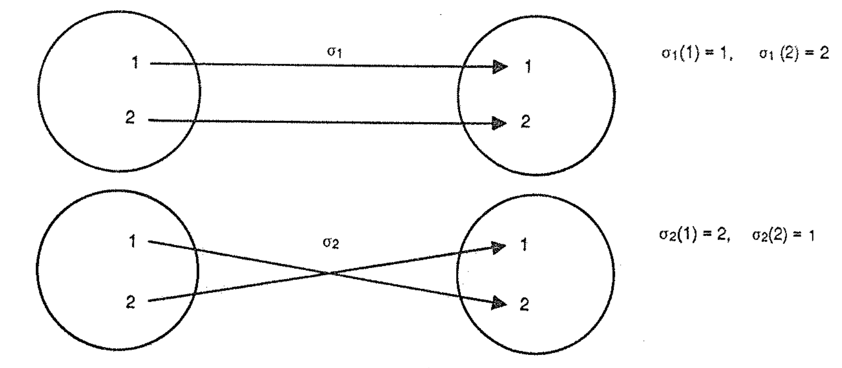
1. PERMUTACIONES

Considérese el conjunto de números $\{1, 2, \ldots, n\}$. Una *permutación* de este conjunto no es más que una determinada ordenación de sus elementos. Por ejemplo, si el conjunto es $\{1, 2\}$, se tienen dos diferentes permutaciones; a saber $(1, 2)$ y $(2, 1)$.

Con objeto de ganar precisión y una adecuada notación para este concepto, establézcase formalmente la definición de *permutación*.

DEFINICIÓN 1.1 Se llama permutación del conjunto $S = \{1, 2, \ldots, n\}$ a una función biyectiva* $\sigma: S \to S$.

EJEMPLO 1 Por ejemplo, si $S = \{1, 2\}$ existen dos funciones biyectivas σ_1 y σ_2 de S en sí mismo



La manera como se escribirá la permutación $\sigma: S \to S$ es

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}$$

o bien, abreviadamente

$$(\sigma(1), \ \sigma(2), \ldots, \ \sigma(n))$$

Así, en el ejemplo anterior la primera permutación es $(1, 2)$ y la segunda $(2, 1)$.

EJEMPLO 2 Para el conjunto $S = \{1, 2, 3\}$ se tienen las siguientes 6 permutaciones

$(1, 2, 3)$

$(1, 3, 2)$

$(2, 1, 3)$

$(2, 3, 1)$

$(3, 1, 2)$

$(3, 2, 1)$

Véanse cuáles son las permutaciones del conjunto de cuatro elementos $S = \{1, 2, 3, 4\}$.

*Es decir, inyectiva (uno a uno) y sobreyectiva, aunque en este caso (el conjunto S es finito) cualquiera de estas dos propiedades implica a la otra.

Si se escribe $\sigma(1) = 1$, entonces $\sigma(2)$ tiene sólo 3 posibilidades 2, 3 o 4 (pues la función σ es inyectiva). Analícense las tres alternativas por separado.

a) Si $\sigma(2) = 2$, entonces $\sigma(3)$ tiene sólo 2 posibilidades: 3 o 4.

a.1) Si $\sigma(3) = 3$, necesariamente $\sigma(4) = 4$, entonces una permutación es:

$$(1, 2, 3, 4)$$

a.2) Si $\sigma(3) = 4$, necesariamente, $\sigma(4) = 3$, entonces otra permutación es

$$(1, 2, 4, 3)$$

b) Si $\sigma(2) = 3$, entonces $\sigma(3)$ puede ser 2 o 4.

b.1) Si $\sigma(3) = 2$, necesariamente $\sigma(4) = 4$, y se tiene entonces la permutación

$$(1, 3, 2, 4)$$

b.2) Si $\sigma(3) = 4$, necesariamente $\sigma(4) = 2$, y se tiene así la permutación

$$(1, 3, 4, 2)$$

c) Si $\sigma(2) = 4$, entonces $\sigma(3)$ puede ser 2 o 3

c.1) Si $\sigma(3) = 2$, necesariamente $\sigma(4) = 3$, y se tiene la permutación en este caso como

$$(1, 4, 2, 3)$$

c.2) Si $\sigma(3) = 3$, necesariamente $\sigma(4) = 2$ y se tiene entonces la permutación

$$(1, 4, 3, 2)$$

En conclusión, si $\sigma(1) = 1$, se tienen las siguientes 6 permutaciones

(1, 2, 3, 4)
(1, 2, 4, 3)
(1, 3, 2, 4)
(1, 3, 4, 2)
(1, 4, 2, 3)
(1, 4, 3, 2)

Este mismo tipo de razonamiento se puede repetir en el caso que se escoja $\sigma(1) = 2$, $\sigma(1) = 3$ o $\sigma(1) = 4$.

Esquemáticamente se puede representar este razonamiento por medio de la siguiente tabla. Las cuatro diferentes columnas indican las respectivas posibles posiciones de los cuatro elementos de S en la permutación. Una vez que se ha decidido sobre la primera columna (esto es, se ha dado el valor de $\sigma(1)$), se tienen sólo 3 posibilidades para la siguiente. Decidido ya el valor de esta columna (esto es, dado el valor a $\sigma(2)$), se tienen sólo dos posibilidades para la siguiente. Dado el valor a ésta (dando $\sigma(3)$) sólo resta escribir el número que falta en la última columna ($\sigma(4)$).

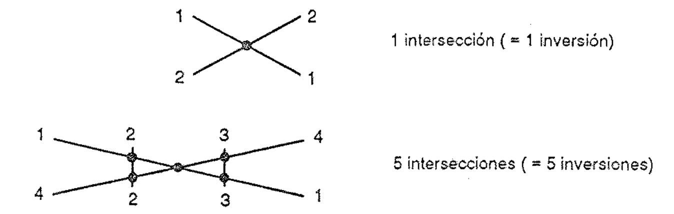
$\sigma(1)$	$\sigma(2)$	$\sigma(3)$	$\sigma(4)$
1	2	3	4
		4	3
	3	2	4
		4	2
	4	2	3
		3	2
2	1	3	4
		4	3
	3	1	4
		4	1
	4	1	3
		3	1
3	1	2	4
		4	2
	2	1	4
		4	1
	4	1	2
		2	1
4	1	2	3
		3	2
	2	1	3
		3	1
	3	1	2
		2	1

Se tienen, entonces, las 24 permutaciones de $\{1, 2, 3, 4\}$ siguientes:

(1, 2, 3, 4)	(2, 1, 3, 4)	(3, 1, 2, 4)	(4, 1, 2, 3)
(1, 2, 4, 3)	(2, 1, 4, 3)	(3, 1, 4, 2)	(4, 1, 3, 2)
(1, 3, 2, 4)	(2, 3, 1, 4)	(3, 2, 1, 4)	(4, 2, 1, 3)
(1, 3, 4, 2)	(2, 3, 4, 1)	(3, 2, 4, 1)	(4, 2, 3, 1)
(1, 4, 2, 3)	(2, 4, 1, 3)	(3, 4, 1, 2)	(4, 3, 1, 2)
(1, 4, 3, 2)	(2, 4, 3, 1)	(3, 4, 2, 1)	(4, 3, 2, 1)

ocurrieron en los trazos realizados (cuidando que no existan intersecciones de 3 o más trazos a la vez). Éste es el número de inversiones de la permutación.*

EJEMPLO 6 Para los dos ejemplos anteriores



DEFINICIÓN 1.3 Sea σ una permutación del conjunto $\{1, 2, \ldots, n\}$. Sea k_σ = número de inversiones de σ. Si k_σ es par, se dice que σ es una *permutación par*; caso contrario, se dice que es una *permutación impar*. Se define el signo de la permutación σ, denotado por *sgn* σ como

$$sgn\ \sigma = (-1)^{k_\sigma}$$

Por lo tanto, si σ es una *permutación* par, *sgn* $\sigma = 1$, caso contrario, si σ es una permutación impar, *sgn* $\sigma = -1$.

EJEMPLO 7 Considere, por ejemplo, las permutaciones del conjunto $\{1, 2, 3\}$. En la siguiente tabla se encuentra la información correspondiente sobre la paridad de estas permutaciones

PERMUTACIÓN	NÚMERO DE INVERSIONES	TIPO DE PERMUTACIÓN	SIGNO
(1, 2, 3)	0	par	1
(2, 3, 1)	2	par	1
(3, 1, 2)	2	par	1
(1, 3, 2)	1	impar	-1
(2, 1, 3)	1	impar	-1
(3, 2, 1)	3	impar	-1

Se quiere ver ahora cómo está relacionado el signo de una composición de permutaciones con el signo de cada una de las permutaciones componentes. Véase primero un ejemplo.

EJEMPLO 8 Sean σ, π las permutaciones

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix} \qquad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 4 & 2 \end{pmatrix}$$

*En efecto: Si $i < j$, considere los trazos que unen $\sigma(i)$ con él mismo y $\sigma(j)$ con él mismo. Estos trazos se intersectan si, y sólo si $\sigma(i) > \sigma(j)$, es decir, si existe una inversión.

Se tiene:

número de inversiones en $\sigma = 3$. Entonces *sgn* $\sigma = -1$.

número de inversiones en $\pi = 2$. Entonces *sgn* $\pi = 1$.

Forme las composiciones $\pi \circ \sigma$ y $\sigma \circ \pi$

$$\pi \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \qquad \sigma \circ \pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$$

Se tiene:

número de inversiones en $\pi \circ \sigma = 1$. Entonces *sgn* $\pi \circ \sigma = -1$.

número de inversiones en $\sigma \circ \pi = 5$. Entonces *sgn* $\sigma \circ \pi = -1$.

Observe que tanto *sgn* $\pi \circ \sigma$ y *sgn* $\sigma \circ \pi$ como el producto (*sgn* σ)(*sgn* π) valen -1. Esto no es una coincidencia. El siguiente teorema nos dice que es un hecho general:

TEOREMA 1.1 Sea σ y π dos permutaciones del conjunto $\{1, 2, \ldots, n\}$. Entonces *sgn* $\sigma \circ \pi$ = (*sgn* σ)(*sgn* π) = *sgn* $\pi \circ \sigma$

DEMOSTRACIÓN Se demostrará que *sgn* $\sigma \circ \pi$ = (*sgn* σ)(*sgn* π). Se debe de contar el número de inversiones que ocurran en π, en σ y en $\sigma \circ \pi$. Tómense los enteros $i, j \in \{1, 2, \ldots, n\}$, tal que $i < j$. La permutación π tendrá una inversión si $\pi(i) > \pi(j)$; similarmente, $\sigma \circ \pi$ tendrá una inversión si $(\sigma \circ \pi)(i) > (\sigma \circ \pi)(j)$. Obsérvese que, partiendo del orden de los enteros $\pi(i)$ y $\pi(j)$ se puede decidir, usando la información del orden de $(\sigma \circ \pi)(i) = \sigma(\pi(i))$ y $(\sigma \circ \pi)(j) = \sigma(\pi(j))$ si σ tiene o no inversión. Es decir, si $\pi(i) < \pi(j)$ (en cuyo caso π no tendrá inversión respecto de i y j), y si $\sigma(\pi(i)) > \sigma(\pi(j))$, σ tendrá una inversión, así como también si $\pi(i) > \pi(j)$ y $\sigma(\pi(i)) < \sigma(\pi(j))$(en cuyo caso π sí tiene inversión respecto de i y j). Se deben entonces de considerar 4 casos

CASO	ORDEN RELATIVO	CONCLUSIÓN
1	$i < j$ $\pi(i) < \pi(j)$ $\sigma(\pi(i)) < \sigma(\pi(j))$	No hay inversión para π No hay inversión para σ No hay inversión para $\sigma \circ \pi$
2	$i < j$ $\pi(i) < \pi(j)$ $\sigma(\pi(i)) > \sigma(\pi(j))$	No hay inversión para π Hay inversión para σ Hay inversión para $\sigma \circ \pi$
3	$i < j$ $\pi(i) > \pi(j)$ $\sigma(\pi(i)) < \sigma(\pi(j))$	Hay inversión para π Hay inversión para σ No hay inversión para $\sigma \circ \pi$
4	$i < j$ $\pi(i) > \pi(j)$ $\sigma(\pi(i)) > \sigma(\pi(j))$	Hay inversión para π No hay inversión para σ Hay inversión para $\sigma \circ \pi$

Llámese $k_{\sigma\circ\pi}$, k_σ, y k_π. a los correspondientes números de inversiones en $\sigma \circ \pi$, σ y π. Obsérvese que cuando hay una inversión en $\sigma \circ \pi$, entonces hay una inversión en σ o (exclusivo) en π (casos 2 y 4 respectivamente). Por lo tanto,

$$k_{\sigma\circ\pi} = k_\sigma + k_\pi - 2r$$

donde r es el número de veces en los que hay inversiones en σ y en π simultáneamente (caso 3). Entonces

$$sgn\ \sigma\circ\pi = (-1)^{k_{\sigma\circ\pi}} = (-1)^{k_\sigma + k_\pi - 2r} = (-1)^{k_\sigma}(-1)^{k_\pi}(-1)^{-2r}$$
$$= (-1)^{k_\sigma}(-1)^{k_\pi} = (sgn\ \sigma)(sgn\ \pi).$$

Q.E.D.

COROLARIO

> Sea σ una *permutación* del conjunto $\{1, 2, \ldots, n\}$. Entonces
>
> $$sgn\ \sigma = sgn\ \sigma^{-1}$$

DEMOSTRACIÓN Como $\sigma \circ \sigma^{-1} = id$ se tiene, $sgn\ \ \sigma\circ\sigma^{-1} = (sgn\ \sigma)(sgn\ \sigma^{-1}) = 1$ (pues $sgn\ id = (-1)^0 = 1$). Entonces $sgn\ \sigma = sgn\ \sigma^{-1}$.

Q.E.D.

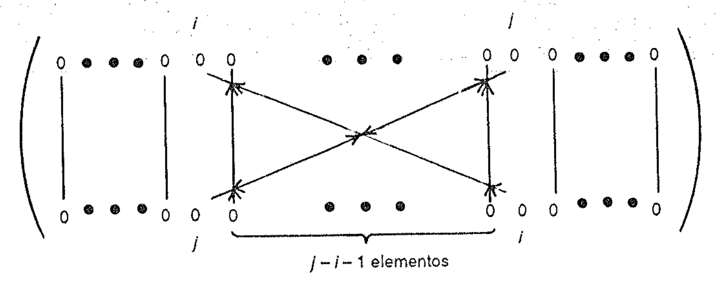
DEFINICIÓN 1.4 Se dice que la permutación σ del conjunto $\{1, 2, \ldots, n\}$ es una *transposición* de los enteros $i, j \in \{1, 2, \ldots, n\}$ si $\sigma(i) = j$, $\sigma(j) = i$ y $\sigma(k) = k$ para $k \neq i,\ j$.

Una transposición es, pues, un tipo muy simple de permutación en la cual sólo es alterada la posición de 2 números i y j: el número i pasa a ocupar la posición de j y viceversa.

Una transposición se ve entonces como

$$\begin{pmatrix} 1 & 2 & \ldots & i & \ldots & j & \ldots & n \\ 1 & 2 & \ldots & j & \ldots & i & \ldots & n \end{pmatrix}$$

Cuéntese el número de inversiones que tiene una transposición de los enteros i y j. Hágase con la regla de "contar intersecciones de trazos". La situación se ve como



Se observa, entonces, que existirán $2(j-i-1) + 1$ intersecciones entre los trazos efectuados. Éste es un número impar. Por lo tanto, se ha probado el siguiente teorema:

TEOREMA 1.2

> Una transposición es una permutación impar.

Por ejemplo, la transposición de los enteros 3 y 9 del conjunto $\{1, 2, \ldots, 10\}$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 1 & 2 & 9 & 4 & 5 & 6 & 7 & 8 & 3 & 10 \end{pmatrix}$$

tiene $2(j-i-1)+1 = 2(9-3-1)+1 = 11$ inversiones, lo cual se puede verificar directamente.

COROLARIO

> Sea σ una permutación cualquiera del conjunto $\{1, 2, \ldots, n\}$. Sea τ una transposición. Así,
>
> $$sgn^{\tau\circ\sigma} = -sgn\ \sigma = sgn^{\sigma\circ\tau}$$

DEMOSTRACIÓN Se sigue inmediatamente de los teoremas 1.1. y 1.2.

Q.E.D.

El resultado anterior dice que si en una permutación se intercambia la posición de dos de sus elementos, este hecho se reflejará en un cambio de signo de la permutación; es decir, la permutación resultante tendrá signo contrario al de la permutación original.

EJERCICIOS (SECCIÓN 1, CAPÍTULO 2)

1. ¿Cuántas permutaciones σ del conjunto $\{1, 2, 3, 4\}$ son tales que $\sigma(1) = 4$ y $\sigma(2) = 3$? Haga una lista de ellas.
2. ¿Cuántas permutaciones σ del conjunto $\{1, 2, 3, 4, 5, 6\}$ son tales que $\sigma(1) = 6$, $\sigma(2) = 5$, $\sigma(3) = 4$? Haga una lista de ellas.
3. Para las siguientes permutaciones σ y π obtenga $\sigma\circ\pi$, $\pi\circ\sigma$, $\sigma\circ\sigma$, $\pi\circ\pi$

a) $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ $\qquad \pi = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$

b) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}$ $\quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$

c) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$ $\quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$

d) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 1 & 4 & 3 \end{pmatrix}$ $\quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 3 & 5 \end{pmatrix}$

e) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 3 & 5 \end{pmatrix}$ $\quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 3 & 2 & 1 \end{pmatrix}$

f) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix}$ $\quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 7 & 1 & 3 & 4 & 2 & 5 \end{pmatrix}$

4. Para las permutaciones del ejercicio anterior obtenga σ^{-1}, π^{-1}, $(\sigma \circ \pi)^{-1}$, $\pi^{-1} \circ \sigma^{-1}$. Compare la permutación $(\sigma \circ \pi)^{-1}$ con $\pi^{-1} \circ \sigma^{-1}$.
5. Sean σ y π dos permutaciones del conjunto $(1, 2, \ldots, n)$. Demuestre que $(\sigma \circ \pi)^{-1} = \pi^{-1} \circ \sigma^{-1}$.
6. Determine el número de inversiones para cada una de las permutaciones σ, π, $\sigma \circ \pi$, $\pi \circ \sigma$, $\sigma \circ \sigma$, $\pi \circ \pi$ del ejercicio 3. Obtenga el signo de cada una de estas permutaciones y verifique que se satisface el teorema 1.1.
7. Establezca el número de inversiones para cada una de las permutaciones σ^{-1}, π^{-1}, $(\sigma \circ \pi)^{-1}$ del ejercicio 4. Verifique que se satisface el corolario del teorema 1.1.
8. Considere la siguiente permutación del conjunto $\{1, 2, \ldots, n\}$:

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ n & n-1 & \ldots & 1 \end{pmatrix}$$

a) Determine el número de inversiones de σ.

b) Obtenga el signo de esta permutación.

c) obtenga σ^{-1}.

9. Considere las siguientes transposiciones σ. Determine el número de inversiones en cada una de ellas y verifique que se satisface el teorema 1.2.

Obtenga σ^{-1} en cada caso.

a) $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$

b) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$

c) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 2 & 3 & 4 & 5 & 1 \end{pmatrix}$

d) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 1 & 11 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 2 & 12 \end{pmatrix}$

10. Sea σ una transposición del conjunto $\{1, 2, \ldots, n\}$. Obtenga σ^{-1}.

② 11. Dada la permutación σ del conjunto $\{1, 2, \ldots, n\}$, defina la matriz asociada a esta permutación, denotada por A_σ, como la matriz cuadrada de orden n, cuyos elementos a_{ij} estás dados por

$$a_{ij} = \begin{cases} 1 \text{ si } j = \sigma(i) \\ 0 \text{ si } j \neq \sigma(i) \end{cases}$$

Por ejemplo, si

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$$

se tiene

$$A_\sigma = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

a) Demuestre que en cada línea de la matriz A_σ se tiene exactamente un uno y ceros en las posiciones restantes y que esta propiedad también vale para las columnas de esta matriz.

b) Sean σ y π las permutaciones

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \qquad \pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

b.1) Halle las matrices A_σ y A_π.

b.2) Encuentre la matriz $A_{\sigma \circ \pi}$.

b.3) Verifique que $A_{\sigma \circ \pi} = A_\pi A_\sigma$.

c) Compruebe que, en general, si A_σ y A_π son las matrices asociadas a las permutaciones σ y π del conjunto $\{1, 2, \ldots, n\}$, entonces $A_{\sigma \circ \pi} = A_\pi A_\sigma$.

d) Sea σ la permutación

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

d.1) Halle A_σ.

d.2) Determine $A_{\sigma^{-1}}$.

d.3) Verifique que $A_\sigma A_{\sigma^{-1}} = A_{\sigma^{-1}} A_\sigma = I_4$.

c) Demuestre que, en general, si A_σ es la matriz asociada a la permutación σ del conjunto $\{1, 2, \ldots, n\}$, entonces $A_\sigma A_{\sigma^{-1}} = A_{\sigma^{-1}} A_\sigma = I_n$.

12. El objetivo de este ejercicio es probar el siguiente resultado general sobre permutaciones: toda permutación $\sigma : S \to S$ del conjunto $S = \{1, 2, \ldots, n\}$ se puede escribir como una composición de transposiciones de elementos de S. Es decir, dada la permutación $\sigma : S \to S$, existen transposiciones $\tau_1, \tau_2, \ldots, \tau_k : S \to S$ tales que $\sigma = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_k$. Esta expresión no es única, pero sí lo es la paridad de k. (Observe entonces que, bajo esta perspectiva, se puede definir una permutación par —impar— como una que se presenta como composición de un número par —impar— de transposiciones). La demostración de este hecho se hace por inducción sobre n. Observe que el resultado es obvio para $n = 2$. Suponga entonces válido el resultado para $n - 1$ y pruébelo para n, siguiendo los incisos siguientes:

a) Sea $\sigma : S \to S$, $S = \{1, 2, \ldots, n\}$ una permutación cualquiera. Si $\sigma(n) = n$. Asocie (restrinja) a σ una permutación $\sigma^* : S^* \to S^*$, $S^* = \{1, 2, \ldots, n-1\}$ definida como $\sigma^*(i) = \sigma(i)$, $i = 1, 2, \ldots, n-1$. Aplique la hipótesis de inducción a σ^* para obtener que $\sigma^* = \tau_1^* \circ \tau_2^* \circ \ldots \circ \tau_k^*$, en donde τ_i^* son transposiciones del conjunto S^*. Extienda las transposiciones τ_i^* a transposiciones τ_i del conjunto S y concluya, finalmente, que $\sigma = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_k$, como se quería.

b) Suponga ahora que $\sigma(n) \neq n$ (y entonces $\sigma(n) < n$). Sea τ la transposición que cambia de posición a los elementos $\sigma(n)$ y n en σ. Observe, entonces que $\tau \circ \sigma$ es una permutación de S en la que $(\tau \circ \sigma)(n) = n$. Concluya, del inciso anterior, que existen transposiciones $\tau_1, \tau_2, \ldots, \tau_k$ tales que $\tau \circ \sigma = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_k$. Demuestre por lo tanto que $\sigma = \tau \circ \tau_1 \circ \tau_2 \circ \ldots \circ \tau_k$, quedando así probado lo que se quería.

c) Defina el número p_n como

$$p_n = \prod_{n \geq i > j \geq 1} (i - j) \qquad i, j \in S$$

Observe que p_n es positivo. Dada la permutación $\sigma : S \to S$, defina el número $\sigma(p_n)$ como

$$\sigma(p_n) = \prod_{n \geq i > j \geq 1} (\sigma(i) - \sigma(j))$$

c.1) Compruebe que $\sigma(p_n) = \pm p_n$.

c.2) Demuestre que si τ es una transposición, entonces $(\tau \circ \sigma)(p_n) = -\sigma(p_n)$.

c.3) Escriba $\sigma = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_k$. Demuestre que $(-1)^k = \sigma(p_n) / p_n$.

c.4) Concluya entonces que la paridad de k depende solamente de la permutación σ y no de la manera de expresarla como composición de transposiciones.

2. DEFINICIÓN Y PROPIEDADES

Una vez establecidas las ideas y resultados que sobre permutaciones se estudiaron en la sección anterior, se está en condiciones de dar una definición satisfactoria de la función determinante.

Antes que nada, se debe decir que el determinante es una *función*, cuyo dominio es el conjunto de matrices cuadradas, que se denotará por M_n (n es el orden de las matrices) y cuyo codominio son los números reales. Se denotará esta función por *det*. Entonces det: $M_n \to R$. Sólo falta entonces dar la *regla* de asociación que define a *det*, es decir, falta dar una *fórmula* para el determinante de una matriz cuadrada. Eso es precisamente lo que se hará ahora.

2.1. DEFINICIÓN DE DETERMINANTE

Sea A la matriz cuadrada de orden n siguiente:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

La *determinante* de A, denotado por $det\, A$, está definida por la fórmula

$$det\, A = \sum_{\sigma} (sgn\ \sigma) a_{1\,\sigma(1)}\, a_{2\,\sigma(2)} \ldots a_{n\,\sigma(n)}$$

donde la suma se hace sobre todas las (n!) permutaciones del conjunto $\{1, 2, \ldots, n\}$

En algunas ocasiones, con objeto de ganar claridad en la estructura de esta fórmula y en las propiedades que se estudiarán en esta sección, se escribirá esta fórmula como

$$det\, A = \sum_{\sigma} sgn \begin{pmatrix} 1 & 2 & & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix} a_{1\,\sigma(1)}\, a_{2\,\sigma(2)} \ldots a_{n\,\sigma(n)}$$

Para hallar el determinante de la matriz A de orden n se debe pues considerar cada una de las permutaciones σ del conjunto $\{1, 2, \ldots, n\}$, dígase $(\sigma(1), \sigma(2), \ldots, \sigma(n))$; con ella, se forma el producto

$$a_{1\,\sigma(1)}\, a_{2\,\sigma(2)} \ldots a_{n\,\sigma(n)}$$

(que se llamará *producto elemental*) el cual está formado por n factores cada uno de los cuales pertenece a una determinada línea y una determinada columna de la matriz A, y en el cual *no* hay dos factores que provengan de la misma línea y/o de la misma columna (**Nota**: el hecho de que no hay dos factores de la misma línea es evidente de la fórmula misma; el hecho de que no hay dos factores de la misma columna se debe a la inyectividad de la permutación σ).

Una vez formado este producto elemental, multiplíquelo por el signo de la permutación σ que se está considerando y finalmente sume todos estos "productos elementales con signo", obteniendo así el determinante de la matriz A.

EJEMPLO 1 Por ejemplo, considere la matriz de orden 2

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Se tiene

PERMUTACIONES DEL CONJUNTO $\{1, 2\}$	$sgn\ \sigma$	PRODUCTO ELEMENTAL $a_{1\,\sigma(1)}a_{2\,\sigma(2)}$	$(sgn\ \sigma)\ a_{1\,\sigma(1)}a_{2\,\sigma(2)}$
(1, 2)	1	$a_{11}\,a_{22}$	$a_{11}\,a_{22}$
(2, 1)	-1	$a_{12}\,a_{21}$	$-a_{12}\,a_{21}$

SUMA $= a_{11}a_{12} - a_{12}a_{21} = det\ A$

o sea,

$$det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

EJEMPLO 2 Para la matriz de orden 3

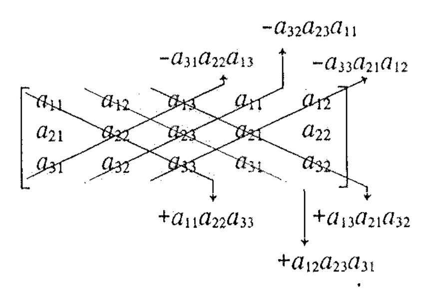
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Se tiene

PERMUTACIONES DEL CONJUNTO $\{1, 2, 3\}$	$sgn\ \sigma$	PRODUCTO ELEMENTAL $a_{1\,\sigma(1)}a_{2\,\sigma(2)}a_{3\,\sigma(3)}$	$(sgn\ \sigma)$ $a_{1\,\sigma(1)}a_{2\,\sigma(2)}a_{3\,\sigma(3)}$
(1, 2, 3)	1	$a_{11}\,a_{22}\,a_{33}$	$a_{11}\,a_{12}\,a_{13}$
(2, 3, 1)	1	$a_{12}\,a_{23}\,a_{31}$	$a_{12}\,a_{22}\,a_{31}$
(3, 1, 2)	1	$a_{13}\,a_{21}\,a_{32}$	$a_{13}\,a_{21}\,a_{32}$
(1, 3, 2)	-1	$a_{11}\,a_{23}\,a_{32}$	$-a_{11}\,a_{23}\,a_{32}$
(2, 1, 3)	-1	$a_{12}\,a_{21}\,a_{33}$	$-a_{12}\,a_{21}\,a_{32}$
(3, 2, 1)	-1	$a_{13}\,a_{22}\,a_{31}$	$-a_{13}\,a_{22}\,a_{31}$

$$\begin{aligned} \textbf{Suma} &= a_{11}a_{12}a_{13} + a_{12}a_{23}a_{31} \\ &+ a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} \\ &- a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \\ &= det\ A \end{aligned}$$

Para el caso particular de matrices de orden 3, se tiene la siguiente regla que permite obtener la fórmula del determinante: reescriba en la parte derecha de la matriz fuera de ella sus dos primeras columnas, quedando de esta manera tres cuadrados sobrepuestos (en un rectángulo de 3 por 5) con "6 diagonales" como en la figura. Forme los productos de los elementos en cada diagonal. Si la diagonal es "decreciente", el producto es positivo, caso contrario, es negativo. La suma de estos productos con signo es el determinante de la matriz.



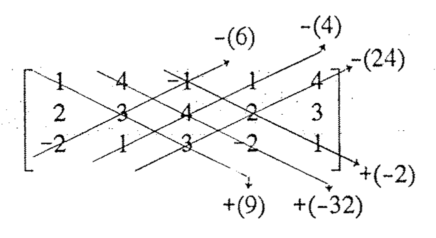
EJEMPLO 3 Véanse un par de ejemplos.

$$det \begin{bmatrix} 2 & 3 \\ 5 & 4 \end{bmatrix} = (2)(4) - (5)(3) = -7$$

La determinante de la matriz

$$A = \begin{bmatrix} 1 & 4 & -1 \\ 2 & 3 & 4 \\ -2 & 1 & 3 \end{bmatrix}$$

se halla según la regla descrita anteriormente



Entonces,

$$det\ A = (9) + (-32) + (-2) - (6) - (4) - (24) = -59$$

2.2. PROPIEDADES

De la definición de determinante establecida en la subsección anterior se ve que para calcular el determinante de una matriz de orden n, se tendrán que realizar $n!$ productos elementales, multiplicarlos por el signo de la permutación correspondiente y luego hacer la suma de todos estos resultados. Ya para una matriz de orden 5, son $5! = 120$ operaciones individuales que se tienen que realizar para luego sumar los resultados. Para una matriz de orden 6 son $6! = 720$ operaciones . . .

No es la intención invitar al lector a calcular determinantes. Más bien se quiere centrar el esfuerzo en estudiar el comportamiento de esta función y algunas de sus propiedades. Resultará que, por añadidura, se obtendrán también alternativas más cortas (mucho más cortas) para evaluar algunos determinantes numéricos.

Sea A una matriz de orden n de elementos a_{ij}, es decir $A = (a_{ij})_{i,j=1,\dots,n}$. Se llama *transpuesta* de A, denotada por A^t, a la matriz $A^t = (a_{ji})_{i,j=1,\dots,n}$.

La transpuesta de una matriz A se obtiene, entonces, intercambiando las líneas por las columnas de la matriz A.

EJEMPLO 4 Por ejemplo, si

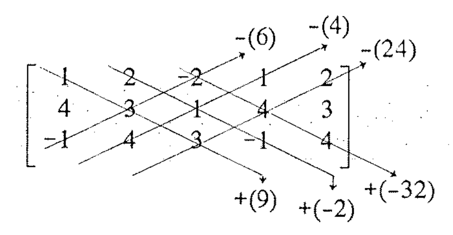
$$A = \begin{bmatrix} 1 & 4 & -1 \\ 2 & 3 & 4 \\ -2 & 1 & 3 \end{bmatrix}$$

se tiene

$$A^t = \begin{bmatrix} 1 & 2 & -2 \\ 4 & 3 & 1 \\ -1 & 4 & 3 \end{bmatrix} \begin{matrix} \leftarrow \text{columnas de } A = \\ \leftarrow \text{líneas de } A^t \end{matrix}$$

↑ ↑ ↑
líneas de A =
columnas de A^t

Ya se había calculado $\det A$ (= −59). Calcúlese $\det A^t$.



Entonces

$$\det A^t = (9) + (-2) + (-32) - (6) - (4) - (24) = -59 = \det A$$

Ésta es una propiedad general que tienen las determinantes que se enunciará y demostrará enseguida:

TEOREMA 2.1 Sea A una matriz de orden n. Entonces

$$\det A = \det A^t$$

DEMOSTRACIÓN Sea $B = (b_{ij})_{i,j=1,\dots,n}$ la matriz transpuesta de $A = (A_{ij})_{i,j=1,\dots,n}$. Es decir, $b_{ij} = a_{ji}$, $i, j = 1, 2, \dots, n$. Entonces la definición de *det*

$$\det A^t = \det B = \sum_{\pi} (sgn\,\pi)\, b_{1\pi(1)} b_{2\pi(2)} \cdots b_{n\pi(n)}$$

Sea $j = \pi(i)$. Entonces $i = \pi^{-1}(j)$ y se puede escribir el elemento $b_{i\pi(i)}$ como $b_{\pi^{-1}(j)j}$. Obsérvese que

$$b_{1\,\pi(1)} b_{2\,\pi(2)} b_{n\,\pi(n)} = b_{\pi^{-1}(1)1} b_{\pi^{-1}(2)2} \cdots b_{\pi^{-1}(n)n}$$

(se han reacomodado los factores de tal modo que queden ordenados atendiendo al segundo subíndice).

Se sabe que $sgn\,\pi^{-1} = sgn\,\pi$ (corolario del teorema 1.1). Siendo así,

$$\det B = \sum (sgn\,\pi^{-1}) a_{1\,\pi^{-1}(1)} a_{2\,\pi^{-1}(2)} \cdots a_{n\,\pi^{-1}(n)}$$

o sea,

$$\det A^t = \sum (sgn\,\pi^{-1}) a_{1\,\pi^{-1}(1)} a_{2\,\pi^{-1}(2)} \cdots a_{n\,\pi^{-1}(n)}$$

Ahora, cada permutación σ del conjunto $\{1, 2, \dots, n\}$ es la inversa de alguna permutación de este conjunto. Es decir, si se consideran las permutaciones σ tales que $\sigma = \pi^{-1}$ para alguna permutación π se están considerando *todas* las permutaciones del conjunto $\{1, 2, \dots, n\}$. Entonces, la fórmula anterior se puede escribir como

$$\det A^t = \sum_{\sigma} (sgn\,\sigma) a_{1\,\sigma(1)} a_{2\,\sigma(2)} \cdots a_{n\,\sigma(n)}$$

que es precisamente la fórmula que define a $\det A$.

Q.E.D.

Se dice que la matriz $A = (a_{ij})_{i,j=1,\dots,n}$ es una *matriz triangular superior* (*triangular inferior*) si $a_{ij} = 0$ para $i > j$ ($a_{ij} = 0$ si $i < j$, respectivamente).

El aspecto que tiene una matriz triangular superior es:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & a_{22} & & a_{2n} \\ & 0 & \ddots & \\ & & & a_{nn} \end{bmatrix}$$

donde el cero indica que la parte de la matriz en donde se encuentra tiene todos sus elementos iguales a cero.

Similarmente, el aspecto de una matriz triangular inferior es:

$$\begin{bmatrix} a_{11} & & 0 & \\ a_{21} & a_{22} & & \\ & \dots & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

El siguiente teorema dice que es particularmente sencillo el cálculo del determinante de una matriz en forma triangular:

TEOREMA 2.2

Sea A una matriz triangular (superior o inferior). Entonces

$$\det A = a_{11}\, a_{22} \dots a_{nn}$$

DEMOSTRACIÓN Obsérvese que la transpuesta de una matriz triangular superior es una matriz triangular inferior. Por lo tanto, según el teorema 2.1, es suficiente probar el resultado para una matriz triangular superior.

Sea σ una permutación del conjunto $\{1, 2, \dots, n\}$. Si para alguna $j \in \{1, 2, \dots, n\}$ acontece que $\sigma(j) \neq j$, entonces debe existir $i \in \{1, 2, \dots, n\}$ tal que $i > \sigma(i)$. En tal caso $a_{i\sigma(i)} = 0$. Siendo así, el único producto elemental que no es cero (que no tiene un factor igual a cero) es aquel en el cual $\sigma(i) = i$ para toda i. Se trata, pues, de la permutación identidad para la cual $sgn\ \sigma = 1$. Entonces

$$\det A = \sum_{\sigma} (sgn\ \sigma) a_{1\,\sigma(1)} \dots a_{n\,\sigma(n)} = a_{11}\, a_{22} \dots a_{nn}$$

Q.E.D.

Por ejemplo:

$$\det \begin{bmatrix} 2 & 5 & 10 \\ 0 & 4 & 9 \\ 0 & 0 & 3 \end{bmatrix} = \det \begin{bmatrix} 2 & 9 & 27 \\ 0 & 4 & 35 \\ 0 & 0 & 3 \end{bmatrix} = \det \begin{bmatrix} 2 & 0 & 0 \\ 59 & 4 & 0 \\ 5 & -20 & 3 \end{bmatrix} = (2)(4)(3) = 24$$

Un tipo especial de matriz triangular es la llamada matriz diagonal: se dice, que la matriz $A = (a_{ij})_{i,j=1,2,\dots,n}$ es una *matriz diagonal* si $a_{ij} = 0$ si $i \neq j$. El aspecto que tiene una matriz diagonal es

$$\begin{bmatrix} a_{11} & & 0 & \\ & a_{22} & & \\ 0 & & \ddots & \\ & & & a_{nn} \end{bmatrix}$$

Se tienen los siguientes corolarios:

COROLARIO 1

A es una matriz diagonal, entonces

$$\det A = a_{11}\, a_{22} \dots a_{nn}$$

DEMOSTRACIÓN Obvio, del teorema anterior.

Q.E.D.

Un caso más particular se encuentra en el siguiente corolario:

COROLARIO 2

$\det I_n = 1$.

DEMOSTRACIÓN En este caso $a_{11} = a_{22} = \dots = a_{nn} = 1$.

Q.E.D.

El siguiente resultado es un resultado técnico. Servirá para establecer una propiedad muy importante del determinante en la próxima subsección.

TEOREMA 2.3

Sean las matrices $A = (a_{ij})_{i,j=1,\dots,n}$, $B = (b_{ij})_{i,j=1,\dots,n}$, $C = (c_{ij})_{i,j=1,\dots,n}$. Supóngase que para alguna $r \in \{1, 2, \dots, n\}$ se tiene $c_{rj} = a_{rj} + b_{rj}$, mientras que si $i \neq r$, $c_{ij} = a_{ij} = b_{ij}$. Entonces

$$\det C = \det A + \det B$$

DEMOSTRACIÓN Se tiene que

$$\det C = \sum_{\sigma} (sgn\ \sigma) c_{1\,\sigma(1)}\, c_{2\,\sigma(2)} \dots c_{r\,\sigma(r)} \dots c_{n\,\sigma(n)}$$

$$= \sum_{\sigma} (sgn\ \sigma) c_{1\,\sigma(1)}\, c_{2\,\sigma(2)} \dots (a_{r\,\sigma(r)} + b_{r\,\sigma(r)}) \dots c_{n\,\sigma(n)}$$

$$= \sum_\sigma (sgn\ \sigma)c_{1\sigma(1)}c_{2\sigma(2)}\ldots a_{r\sigma(r)}\ldots c_{n\sigma(n)}$$

$$+ \sum_\sigma (sgn\ \sigma)c_{1\sigma(1)}c_{2\sigma(2)}\ldots b_{r\sigma(r)}\ldots c_{n\sigma(n)}$$

$$= \sum_\sigma (sgn\ \sigma)a_{1\sigma(1)}a_{2\sigma(2)}\ldots a_{r\sigma(r)}\ldots a_{n\sigma(n)}$$

$$+ \sum_\sigma (sgn\ \sigma)b_{1\sigma(1)}b_{2\sigma(2)}\ldots b_{r\sigma(r)}\ldots b_{n\sigma(n)}$$

$$= det\,A + det\,B$$

Q.E.D.

EJEMPLO 5 Por ejemplo si

$$A = \begin{bmatrix} 1 & 4 \\ 2 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 4 \\ 4 & 8 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 4 \\ 6 & 5 \end{bmatrix}$$

(la segunda línea de C es la suma de las líneas correspondientes de A y B). Entonces,

$$det\,A = det \begin{bmatrix} 1 & 4 \\ 2 & -3 \end{bmatrix} = -11 \qquad det\,B = det \begin{bmatrix} 1 & 4 \\ 4 & 8 \end{bmatrix} = -8$$

$$det\,C = det \begin{bmatrix} 1 & 4 \\ 6 & 5 \end{bmatrix} = -19 = -11 - 8 = det\,A + det\,B$$

2.3 OPERACIONES ELEMENTALES Y DETERMINANTES

Sea dada la matriz A de orden n. Se encuentra uno interesado en ver cómo se afecta el valor del determinante de A cuando se realiza en ésta alguna operación elemental.

TEOREMA 2.4

> Si A' es la matriz que se obtiene de A intercambiando dos de sus líneas, entonces
>
> $$det\,A' = -\,det\,A$$

DEMOSTRACIÓN Sea $A' = (a'_{ij})_{i,j=1,\ldots,n}$. Supóngase que A' se obtuvo de A intercambiando en ésta las líneas p y q. Entonces si $A = (a_{ij})_{i,j=1,\ldots,n}$ se tiene

$$a'_{ij} = a_{ij} \text{ si } i \neq p, q$$
$$a'_{pj} = a_{qj}$$
$$a'_{qj} = a_{pj}$$

Se tiene por lo tanto,

$$det\,A' = \sum_\sigma (sgn\ \sigma)a'_{1\sigma(1)}\ldots a'_{p\sigma(p)}\ldots a'_{q\sigma(q)}\ldots a'_{n\sigma(n)}$$

$$= \sum_\sigma (sgn\ \sigma)a_{1\sigma(1)}\ldots a_{q\sigma(p)}\ldots a_{p\sigma(q)}\ldots a_{n\sigma(n)}$$

Esta fórmula se parece mucho a la que describe el determinante de la matriz A. Sólo que se tiene un problema: la permutación σ que aparece en $(sgn\ \sigma)$ no corresponde a la permutación que aparece en el segundo subíndice del producto elemental correspondiente.

Para ver más claro el problema se escribe

$$det\,A' = \sum_\sigma sgn \begin{pmatrix} 1 & \ldots & \boxed{p} & \ldots & \boxed{q} & \ldots & n \\ \sigma(1) & \ldots & \boxed{\sigma(p)} & \ldots & \boxed{\sigma(q)} & \ldots & \sigma(n) \end{pmatrix} a_{1\sigma(1)}\ldots a_{q\sigma(p)}\ldots a_{p\sigma(q)}\ldots a_{n\sigma(n)}$$

mientras que la fórmula de $det\,A$ debe ser

$$det\,A' = \sum_\pi sgn \begin{pmatrix} 1 & \ldots & \boxed{p} & \ldots & \boxed{q} & \ldots & n \\ \pi(1) & \ldots & \boxed{\pi(p)} & \ldots & \boxed{\pi(q)} & \ldots & \pi(n) \end{pmatrix} a_{1\pi(1)}\ldots a_{q\pi(q)}\ldots a_{p\pi(p)}\ldots a_{n\pi(n)}$$

Este problema se resuelve con una transposición: Sea τ la transposición

$$\tau = \begin{pmatrix} 1 & \ldots & p & \ldots & q & \ldots & n \\ 1 & \ldots & q & \ldots & p & \ldots & n \end{pmatrix}$$

Entonces $\pi = \sigma \circ \tau$ es la permutación

$$\pi = \begin{pmatrix} 1 & \ldots & p & \ldots & q & \ldots & n \\ \pi(1) & \ldots & \pi(p) & \ldots & \pi(q) & \ldots & \pi(n) \end{pmatrix} = \begin{pmatrix} 1 & \ldots & p & \ldots & q & \ldots & n \\ \sigma(1) & \ldots & \sigma(q) & \ldots & \sigma(p) & \ldots & \sigma(n) \end{pmatrix}$$

Obsérvese que mientras σ recorre todas las permutaciones de $\{1, 2, \ldots, n\}$, asimismo la hará π (sólo que en otro orden). Además, $sgn\ \pi = sgn\ (\sigma \circ \tau) = -sgn\ \sigma$ (corolario del teorema 1.2). Entonces

$$det\,A = \sum_\pi (sgn\ \pi)a_{1\pi(1)}\ldots a_{q\pi(q)}\ldots a_{p\pi(p)}\ldots a_{n\pi(n)}$$

$$= -\sum_\sigma (sgn\ \sigma)a_{1\sigma(1)}\ldots a_{q\sigma(p)}\ldots a_{p\sigma(q)}\ldots a_{n\sigma(n)}$$

$$= -\,det\,A'$$

Q.E.D.

EJEMPLO 6 Por ejemplo:

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

$$\det\begin{bmatrix} c & d \\ a & b \end{bmatrix} = cb - da = -(ad - bc) = -\det\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Un corolario importante del teorema anterior es el siguiente:

COROLARIO

> Si A es una matriz de orden n, que posee dos líneas iguales, entonces $\det A = 0$.

DEMOSTRACIÓN Supóngase que A tiene las líneas p y q iguales. Sea A' la matriz que se obtiene de A intercambiando las líneas p y q. Es claro que $A' = A$. Pero, según el teorema anterior

$$\det A' = \det A = -\det A$$

Entonces $\det A = 0$.

Q.E.D.

TEOREMA 2.5

> Sea A' la matriz que se ha obtenido de A multiplicando una de sus líneas por la constante c. Entonces
>
> $$\det A' = c \det A$$

DEMOSTRACIÓN Sea $A' = (a'_{ij})_{i,j=1,\ldots,n}$, $A = (a_{ij})_{i,j=1,\ldots,n}$. Dígase que la línea r de A fue multiplicada por c para obtener A'. Entonces

$$a'_{ij} = a_{ij} \text{ si } i \neq r$$

$$a'_{rj} = ca_{rj}$$

Entonces

$$\det A' = \sum_{\pi} (\operatorname{sgn} \pi) a'_{1\,\pi(1)} \ldots a'_{r\,\pi(r)} \ldots a'_{n\,\pi(n)}$$

$$= \sum_{\pi} (\operatorname{sgn} \pi) a_{1\,\pi(1)} \ldots (ca_{r\,\pi(r)}) \ldots a_{n\,\pi(n)}$$

$$= c\sum_{\pi} (\operatorname{sgn} \pi) a_{1\,\pi(1)} \ldots a_{r\,\pi(r)} \ldots a_{n\,\pi(n)}$$

$$= c \det A$$

Q.E.D.

EJEMPLO 7 Por ejemplo, si

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 8 \end{bmatrix} \qquad (\det A = -2)$$

y A' es la matriz que se obtiene de A multiplicando su primera línea por 6, es decir,

$$A' = \begin{bmatrix} 6 & 12 \\ 5 & 8 \end{bmatrix}$$

se tiene

$$\det A' = \det\begin{bmatrix} 6 & 12 \\ 5 & 8 \end{bmatrix} = -12 = 6(-2) = 6 \det A$$

TEOREMA 2.6

> Sea A' la matriz que se obtiene de A sustituyendo su r-ésima línea por ella más k veces su s-ésima línea. Entonces,
>
> $$\det A' = \det A$$

DEMOSTRACIÓN Sea $B = (b_{ij})_{i,j=1,\ldots,n}$ la matriz tal que

$$b_{ij} = a_{ij} \text{ si } i \neq r$$

$$b_{rj} = k\,a_{sj}$$

Sea $A' = (a'_{ij})_{i,j=1,\ldots,n}$. Obsérvese que

$$a'_{ij} = a_{ij} = b_{ij} \text{ si } i \neq r$$

$$a'_{rj} = a_{rj} + ka_{sj} = a_{rj} + b_{rj}$$

Se está entonces dentro de las hipótesis del teorema 2.3, del cual se concluye

$$\det A' = \det A + \det B$$

Pero, según el teorema 2.5, $\det B = k \det \tilde{A}$ donde $\tilde{A}$ es una matriz que tiene sus líneas r y s iguales, y por tanto, por el corolario del teorema 2.4, $\det \tilde{A} = 0$; esto es, $\det B = 0$. Entonces $\det A' = \det A$, como se quería.

Q.E.D.

EJEMPLO 8 Por ejemplo, si

$$A = \begin{bmatrix} 1 & 5 \\ -2 & 4 \end{bmatrix}$$

Haga $(L1) \to (L1) + 10(L2)$, obteniendo

$$A' = \begin{bmatrix} -19 & 45 \\ -2 & 4 \end{bmatrix}$$

Se tiene

$$\det A' = \det \begin{bmatrix} -19 & 45 \\ -2 & 4 \end{bmatrix} = 14 = \det \begin{bmatrix} 1 & 5 \\ -2 & 4 \end{bmatrix} = \det A$$

En resumen, en el siguiente cuadro se muestra cómo afectan las operaciones elementales en una matriz el valor de su determinante:

OPERACIÓN ELEMENTAL	EFECTO EN *det A*
Intercambio de líneas en A	$\det A' = -\det A$
Multiplicar una línea por la constante c	$\det A' = c \det A$
Sustituir una línea por ella misma más un múltiplo de otra	$\det A' = \det A$

(A' es la matriz que se obtiene de A haciendo en esta última la operación elemental indicada).

NOTA IMPORTANTE

En el teorema 2.1 se ha probado que el determinante de una matriz cuadrada A es igual al determinante de su matriz transpuesta A^t. Una consecuencia importante de este hecho es que toda propiedad de los determinantes relacionada con las líneas de una matriz sigue siendo válida si en lugar de las líneas de la matriz se consideran sus columnas (pues las columnas de la matriz son las líneas de la matriz transpuesta, y el determinante no se ve afectado por este cambio).

Entonces, las propiedades enunciadas en los teoremas 2.4, 2.5 y 2.6 son igualmente válidas si se cambia la palabra "líneas" por la palabra "columnas". Más precisamente, combinando el teorema 2.1 con los teoremas 2.4, 2.5 y 2.6, se tienen los siguientes resultados.

OPERACIÓN REALIZADA EN *A*	EFECTO EN *det A*
Intercambio de columnas de A	$\det A' = -\det A$
Multiplicar una columna por la letra c	$\det A' = c \det A$
Sustituir una columna por ella misma más un múltiplo de otra	$\det A' = \det A$

(A' es la matriz que se obtiene de A realizando las operaciones indicadas en esta última).

Las ideas anteriores pueden ser usadas para realizar cálculos de determinantes de una matriz haciendo en ella operaciones adecuadas en sus líneas y en sus

columnas. Antes de analizar algunos ejemplos en los que se apliquen estas propiedades, véase un último resultado acerca del comportamiento del determinante.

TEOREMA 2.7 Si la matriz cuadrada A tiene una línea (o columna) de ceros, entonces $\det A = 0$.

DEMOSTRACIÓN Si r es la línea de A que consta sólo de ceros, esto es, $a_{rj} = 0, j = 1, 2, \ldots, n$, basta observar que todos los productos elementales que aparecen en la definición de *det A* tienen siempre como factor un elemento de esta línea. Por tanto todos son cero, y entonces $\det A = 0$. Por la nota anterior, el resultado vale también para columnas.

Q.E.D.

COROLARIO Si A es una matriz no inversible, entonces $\det A = 0$.

DEMOSTRACIÓN Primeramente obsérvese que si A es una matriz no inversible, entonces su forma escalonada reducida tendrá al menos una línea de ceros. En efecto, supóngase, para obtener una contradicción, que la forma escalonada reducida de A no tiene línea alguna de ceros. La única matriz de orden n en forma escalonada reducida sin líneas de ceros es la matriz identidad I_n. Pero, siendo así, esto querría decir que partiendo de A se puede llegar, realizando en ésta operaciones elementales, a la matriz I_n. Es decir, A sería equivalente a I_n. Entonces, por el teorema 4.7 (capítulo 1), A sería una matriz inversible, lo cual es una contradicción.

Al partir entonces de A y haciendo operaciones elementales, se llega a A', su forma escalonada reducida, que es una matriz que tiene al menos una línea de ceros. El valor del determinante de A se irá afectando en cada operación elemental, por un cambio de signo o por la multiplicación de una constante (según lo discutido anteriormente). Pero finalmente se llega a A', cuyo determinante es igual a cero (teorema anterior). Entonces $\det A = 0$.

Q.E.D.

Véanse algunos ejemplos:

EJEMPLO 9 Se quiere calcular

$$\det \begin{bmatrix} 1 & -1 & 2 & 1 \\ -1 & 3 & 2 & -2 \\ -3 & -3 & 1 & 1 \\ 4 & 2 & 1 & 0 \end{bmatrix}$$

Haga operaciones elementales en las líneas de A con objeto de llevarla a la forma escalonada. Se tiene

$$\det\begin{bmatrix} 1 & -1 & 2 & 1 \\ -1 & 3 & 2 & -2 \\ -3 & -3 & 1 & 1 \\ 4 & 2 & 1 & 0 \end{bmatrix} = \det\begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 2 & 4 & -1 \\ 0 & -6 & 7 & 4 \\ 0 & 6 & -7 & -4 \end{bmatrix} = \det\begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 2 & 4 & -1 \\ 0 & -6 & 7 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$(L4) \to (L4) - 4(L1)$
$(L2) \to (L2) + (L1)$ $\quad (L4) \to (L4) + (L3)$
$(L3) \to (L3) + 3(L1)$

Aunque el objetivo era llevar la matriz a la forma escalonada, se ha encontrado ya con una línea de ceros. Por el teorema 2.7 el valor del determinante es cero. Observe que todas las operaciones elementales que se realizaron en la matriz no alteraron el valor del determinante.

EJEMPLO 10 Suponga ahora que se quiere calcular

$$\det\begin{bmatrix} 1 & 2 & 4 & 5 \\ 3 & 1 & 2 & 5 \\ 4 & 5 & -2 & 2 \\ -1 & 4 & 1 & 4 \end{bmatrix}$$

Se expone enseguida la secuencia de operaciones realizadas en la matriz y el teorema sobre determinantes aplicado en cada paso

$$\det\begin{bmatrix} 1 & 2 & 4 & 5 \\ 3 & 1 & 2 & 5 \\ 4 & 5 & -2 & 2 \\ -1 & 4 & 1 & 4 \end{bmatrix} \overset{\text{teo. 2.6}}{=} \det\begin{bmatrix} 1 & 2 & 4 & 5 \\ 0 & -5 & -10 & -10 \\ 0 & -3 & -18 & -18 \\ 0 & 6 & 5 & 9 \end{bmatrix} \overset{\text{teo. 2.5}}{=} (-5)\det\begin{bmatrix} 1 & 2 & 4 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & -3 & -18 & -18 \\ 0 & 6 & 5 & 9 \end{bmatrix} =$$

$(L2) \to (L2) - 3(L1)$
$(L3) \to (L3) - 4(L1)$ $\quad (L2) \to -\frac{1}{5}(L2)$
$(L4) \to (L4) + (L1)$

$$\overset{\text{teo. 2.6}}{=} (-5)\det\begin{bmatrix} 1 & 2 & 4 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & -12 & -12 \\ 0 & 0 & -7 & -3 \end{bmatrix} \overset{\text{teo. 2.5}}{=} (-5)(-12)\det\begin{bmatrix} 1 & 2 & 4 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -7 & -3 \end{bmatrix}$$

$(L3) \to (L3) + 3(L2)$ $\quad (L3) \to -\frac{1}{12}(L3)$
$(L4) \to (L4) - 6(L2)$

$$\overset{\text{teo. 2.6}}{=} (-5)(-12)\det\begin{bmatrix} 1 & 2 & 4 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix} \overset{\text{teo. 2.2}}{=} (-5)(-12)(4) = 240$$

$(L4) \to (L4) + 7(L3)$

Y así, en lugar de las $4! = 24$ operaciones que se hubieran tenido que realizar para hacer el cálculo del determinante, sólo se tuvieron que hacer 8 operaciones en las líneas de la matriz y aplicar las propiedades correspondientes de los determinantes.

EJEMPLO 11 Como último ejemplo, considere la siguiente ecuación:

$$\det\begin{bmatrix} 1 & 1 & 1 \\ x & x_1 & x_2 \\ y & y_1 & y_2 \end{bmatrix} = 0$$

Se afirma que ésta es la ecuación de una recta que pasa por los puntos (x_1, y_1) y (x_2, y_2).

En efecto, observe que se trata de una ecuación lineal en x y en y (tanto x como y aparecen sólo en una línea y en una columna. Al desarrollar el determinante se obtendría una expresión del tipo $Ax + By + C = 0$ para ciertas constantes A, B, C). Además vea que tanto $x = x_1, y = y_1$ como $x = x_2, y = y_2$ satisfacen la ecuación (pues en el primer caso se tendría el determinante de una matriz con su primera y segunda columna iguales, y en el segundo, el determinante de una matriz con su primera y tercera columna iguales; en ambos casos el determinante es cero).

En conclusión, se tiene una ecuación lineal en x y y que se satisface con los valores $x = x_1, y = y_1$ y $x = x_2, y = y_2$. Es pues, la ecuación de una recta que pasa por (x_1, y_1) y (x_2, y_2).

Para finalizar esta sección, se enuncia el siguiente teorema en el que se resumen las principales propiedades de la función determinante estudiadas aquí:

TEOREMA 2.8 Sea A una matriz cuadrada de orden n, dígase $A = (a_{ij})_{i,j=1,\ldots,n}$.

a) $\det A = \det A^t$.
b) Si A es una matriz triangular (superior o inferior) entonces el determinante de A es igual al producto de los elementos de su diagonal principal.
   b.1) $\det \mathrm{diag}(a_1, a_2, \ldots, a_n) = a_1 a_2 \ldots a_n$.
   b.2) $\det I_n = 1$.
c) Si A' es la matriz que se obtiene de A intercambiando en ésta dos de sus líneas, entonces $\det A' = -\det A$.
d) Si A' es la matriz que resulta de A multiplicando una de sus líneas por la constante k, entonces $\det A' = k \det A$.
e) Si A' es la matriz que se obtiene de A sustituyendo una de sus líneas por ella misma más un múltiplo de otra línea, entonces $\det A' = \det A$.

f) Si la matriz A tiene dos de sus líneas iguales, entonces $det\, A = 0$.
g) Si la matriz A tiene una de sus líneas formada exclusivamente de ceros, entonces $det\, A = 0$.
h) Las propiedades enunciadas en los incisos c), d), e), f) y g) anteriores, siguen siendo válidas si en lugar de la palabra "líneas" se escribe la palabra "columnas".
i) Si la matriz A no es inversible, entonces $det\, A = 0$.

DEMOSTRACIÓN Dispersa en las páginas anteriores.

Q.E.D.

EJERCICIOS (SECCIÓN 2, CAPÍTULO 2)

1. Considere la matriz A de orden 5 con elementos a_{ij}, $i,j = 1, 2, \ldots, 5$. Determine cuáles de los siguientes productos de elementos de la matriz A aparecen en la expresión de $det\, A$ y con qué signo

   1) $a_{11}a_{22}a_{33}a_{44}a_{55}$
   2) $a_{15}a_{23}a_{32}a_{41}a_{54}$
   3) $a_{11}a_{21}a_{32}a_{43}a_{54}$
   4) $a_{15}a_{24}a_{33}a_{42}a_{51}$
   5) $a_{11}a_{32}a_{25}a_{53}a_{44}$
   6) $a_{15}a_{53}a_{42}a_{21}a_{11}$
   7) $a_{51}a_{42}a_{34}a_{23}a_{11}$
   8) $a_{55}a_{45}a_{32}a_{21}a_{14}$

2. Haga una lista de todos los productos elementales que aparecen en la expresión del determinante de una matriz de orden 4 $A = (a_{ij})_{i,\,j=1,\ldots,4}$ que incluyen el factor a_{23} y encuentre su signo.
3. Realice una lista de todos los productos elementales que aparecen en la expresión del determinante de una matriz de orden 5 $A = (a_{ij})_{i,j=1,\ldots,5}$ que incluyen el factor $a_{11}a_{34}$ y determine su signo.
4. Haga una lista de todos los productos elementales que aparecen en la expresión del determinante de una matriz de orden 6 $A = (a_{ij})_{i,j=1,\ldots,6}$ tales que
   a) aparezca el factor $a_{11}a_{22}$ y que tengan signo positivo,
   b) descubra el factor $a_{11}a_{32}a_{43}$ y que tengan signo positivo,
   c) aparezca el factor $a_{21}a_{32}a_{13}$ y que tengan signo negativo,
   d) descubra el factor $a_{11}a_{22}a_{33}a_{44}$ y que tengan signo negativo.
5. Determine para qué valores de i y j el producto $a_{12}a_{2i}a_{35}a_{4j}a_{54}$ es un producto elemental que aparece en la expresión del determinante de una matriz de orden 5 con signo positivo.
6. Determine para qué valores de i, j, k el producto $a_{5i}a_{1k}a_{41}a_{2j}a_{62}a_{34}a_{73}$ es un producto elemental que se halla en la expresión del determinante de una matriz A de orden 7 con signo negativo.
7. Considere la matriz

$$A = \begin{bmatrix} 2a & 1 & 4 \\ 8 & 7 & 5a \\ 3 & a & -2 \end{bmatrix}$$

Sin hacer el cálculo del determinante de esta matriz, encuentre el coeficiente de a^3 en la expresión de $det\, A$.

8. Considere la matriz

$$A = \begin{bmatrix} 2 & -1 & 2x & 3 & 4 \\ x & -2 & 3 & 1 & 4 \\ 5 & 3 & 2 & 3x & 1 \\ 1 & -x & 4 & 2 & 2 \\ 3 & 4 & 8 & 1 & 3x \end{bmatrix}$$

En la expresión de $det\, A$ aparece un término del tipo kx^5. Calcule el valor de k.

9. Considere la matriz

$$A = \begin{bmatrix} 2a & 2 & 3b & 5 \\ -1 & b & 2a & 4 \\ 8 & a & b & 2b \\ -b & -10 & 5 & 3a \end{bmatrix}$$

Determine cuáles de los siguientes términos aparecen en la expresión de $det\, A$. En el caso de que aparezcan, calcule el coeficiente k.

   a) ka^4
   b) kb^4
   c) kb^5
   d) ka^2b^3

10. Calcule el determinante de la matriz de orden n

$$A = \begin{bmatrix} 0 & & & 1 \\ & & 1 & \\ & \cdot^{\cdot^{\cdot}} & & 0 \\ 1 & & & \end{bmatrix}$$

11. Sea $A\sigma$ la matriz asociada a la permutación σ del conjunto $\{1, 2, \ldots, n\}$ (ver ejercicio 11 de la sección anterior). Demuestre que $det\, A\sigma = sgn\sigma$.
12. Al usar la definición de determinante, calcule

$$det \begin{bmatrix} & & & & a_{1n} \\ & 0 & & a_{2,\,n-1} & a_{2n} \\ & & \cdot^{\cdot^{\cdot}} & & \\ a_{n1} & a_{n2} & \cdots & a_{n,\,n-1} & a_{nn} \end{bmatrix}$$

13. Al usar la regla para calcular determinantes de orden 3 dada en la subsección 2.1, evalúe los siguientes determinantes:

a) $det\begin{bmatrix} 1 & 2 & 5 \\ 3 & 4 & -2 \\ 1 & 1 & 4 \end{bmatrix}$ c) $det\begin{bmatrix} 4 & 8 & -1 \\ 3 & 2 & -2 \\ 2 & 2 & 1 \end{bmatrix}$

b) $det\begin{bmatrix} 1 & 2 & 8 \\ -3 & 4 & 2 \\ 1 & -1 & 4 \end{bmatrix}$ d) $det\begin{bmatrix} 1 & 1 & 4 \\ 2 & 8 & -1 \\ 3 & 9 & 3 \end{bmatrix}$

¿Es esta regla válida para matrices de orden 4? Si es así, demuéstrelo. Caso contrario, dé un contraejemplo.

14. En la subsección 2.2 se definió el concepto de *transpuesta* de una matriz cuadrada. Compruebe las siguientes propiedades de la *transpuesta* (A y B son matrices cuadradas del mismo orden)

a) $(A + B)^t = A^t + B^t$ b) $(cA)^t = cA^t$ c) $(AB)^t = B^tA^t$

15. Sea A una matriz de orden n y c un número real. Demuestre que

$$det\,(cA) = c^n\, det\, A$$

(*Sugerencia*: use el teorema 2.5.)

16. Se dice que la matriz cuadrada A es una matriz: (1) *simétrica* si $A = A^t$; (2) *antisimétrica* si $A = -A^t$.

a) Demuestre que si A y B son matrices simétricas del mismo orden, entonces la matriz $A + B$ también es simétrica.

b) Compruebe que si A es una matriz simétrica y c es un número real, entonces la matriz cA también es simétrica.

c) Demuestre la validez de los resultados de los dos incisos anteriores para matrices antisimétricas.

d) Describa la estructura general de las matrices simétricas de orden 2 y de orden 3.

e) Describa la estructura general de las matrices antisimétricas de orden 2 y de orden 3.

f) El siguiente argumento "demuestra" que toda matriz cuadrada es simétrica: Sea A una matriz cuadrada cualquiera. Según el teorema 2.1 se tiene que $det\,A = det\,A^t$. Entonces $A = A^t$ y por tanto A es simétrica. Diga en dónde se encuentra el error de este argumento.

g) Sea A una matriz cuadrada de orden n. Demuestre que la matriz $A_1 = A + A^t$ es simétrica.

h) Sea A una matriz cuadrada de orden n. Pruebe que la matriz $A_2 = A - A^t$ es antisimétrica.

i) Use los incisos b), c), g) y h) para demostrar que toda matriz cuadrada A se puede escribir como la suma de una matriz simétrica y una matriz antisimétrica. Verifique este resultado con las matrices

i.1) $A = \begin{bmatrix} 2 & 8 \\ 9 & 10 \end{bmatrix}$ i.2) $A = \begin{bmatrix} 1 & 3 & -2 \\ 5 & 7 & 8 \\ 4 & 5 & -10 \end{bmatrix}$

j) Compruebe que la matriz cero de orden n es simétrica y antisimétrica a la vez. Más aún, demuestre que ésta es la única matriz con tal propiedad.

k) Demuestre que si A es una matriz antisimétrica de orden n, con n impar, entonces $det\,A = 0$. (*Sugerencia*: use el resultado del inciso anterior.)

l) Por medio de un ejemplo concreto, muestre que la afirmación del inciso anterior es falsa si n no es impar.

17. Suponga que

$$det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = 3$$

Calcule:

a) $det\begin{bmatrix} b & e & f \\ a & b & c \\ g & h & i \end{bmatrix}$ d) $det\begin{bmatrix} c & 4a & b \\ f & 4d & e \\ i & 4g & h \end{bmatrix}$

b) $det\begin{bmatrix} 2a & 2b & 2c \\ g & h & i \\ 3d & 3e & 3f \end{bmatrix}$ e) $det\begin{bmatrix} d+2f & f & e \\ a+2c & c & b \\ g+2i & i & h \end{bmatrix}$

c) $det\begin{bmatrix} a+d & b+e & f+c \\ g & h & i \\ d & e & f \end{bmatrix}$ f) $det\begin{bmatrix} -i & -g & -h \\ f+c & d+a & e+b \\ c & a & b \end{bmatrix}$

18. Para la matriz cuadrada A de orden n se define la matriz Ac como aquella que resulta de cambiar de posición los elementos de A de modo que cada elemento quede en la posición simétrica respecto del centro de la matriz. Por ejemplo:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad Ac = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \qquad Ac = \begin{bmatrix} 16 & 15 & 14 & 13 \\ 12 & 11 & 10 & 9 \\ 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{bmatrix}$$

¿Cómo es el valor del determinante de la matriz Ac respecto del valor del determinante de la matriz A?

② 19. Resuelva la ecuación

a) $det\begin{bmatrix} 1 & x & x^2 & x^3 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix} = 0$ b) $det\begin{bmatrix} x & a & b & c \\ a & x & b & c \\ a & b & x & c \\ a & b & c & x \end{bmatrix} = 0$

(*Sugerencia*: inciso a): observe que el miembro izquierdo de esta expresión es un polinomio de grado 3. El problema es entonces hallar las raíces de este polinomio. Aplique el corolario del teorema 2.4. Una observación similar para el inciso b): el polinomio del miembro izquierdo es ahora de grado 4.)

20. Resuelva la ecuación

$$det\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1-x & 1 & \dots & 1 \\ 1 & 1 & 2-x & \dots & 1 \\ & & \cdot \ \cdot \ \cdot & & \\ 1 & 1 & 1 & \dots & n-x \end{bmatrix} = 0$$

21. Al usar solamente las propiedades de la función determinante descritas en los teoremas 2.4, 2.5, 2.6 y 2.7, calcule los siguientes determinantes:

a) $det\begin{bmatrix} a & b & c \\ b+c & a+c & a+b \\ 1 & 1 & 1 \end{bmatrix}$

d) $det\begin{bmatrix} \operatorname{sen} a & \cos a & \cos(a+d) \\ \operatorname{sen} b & \cos b & \cos(b+d) \\ \operatorname{sen} c & \cos c & \cos(c+d) \end{bmatrix}$

b) $det\begin{bmatrix} (a_1+b_1)^2 & (a_2+b_2)^2 & (a_3+b_3)^2 \\ a_1^2+b_1^2 & a_2^2+b_2^2 & a_3^2+b_3^2 \\ a_1b_1 & a_2b_2 & a_3b_3 \end{bmatrix}$

e) $det\begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}$

c) $det\begin{bmatrix} a & b & c \\ d & e & f \\ ma+nd & mb+ne & mc+nf \end{bmatrix}$

f) $det\begin{bmatrix} 1^2 & 2^2 & 3^2 & 4^2 \\ 2^2 & 3^2 & 4^2 & 5^2 \\ 3^2 & 4^2 & 5^2 & 6^2 \\ 4^2 & 5^2 & 6^2 & 7^2 \end{bmatrix}$

③ 22. Sin hacer el cálculo de los determinantes involucrados, pruebe la validez de las siguientes expresiones:

a) $det\begin{bmatrix} a_1 & b_1 & a_1x+b_1y+c_1 \\ a_2 & b_2 & a_2x+b_2y+c_2 \\ a_3 & b_3 & a_3x+b_3y+c_3 \end{bmatrix} = det\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$

b) $det\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ bc & ac & ab \end{bmatrix} = (b-a)(c-a)(c-b)$

c) $det\begin{bmatrix} 1 & a & a^3 \\ 1 & b & b^3 \\ 1 & c & c^3 \end{bmatrix} = (a+b+c)\, det\begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}$

d) $det\begin{bmatrix} (b+c)^2 & a^2 & a^2 \\ b^2 & (c+a)^2 & b^2 \\ c^2 & c^2 & (a+b)^2 \end{bmatrix} = 2abc(a+b+c)^3$

e) $det\begin{bmatrix} 1+a & 1 & 1 & 1 \\ 1 & 1+b & 1 & 1 \\ 1 & 1 & 1+c & 1 \\ 1 & 1 & 1 & 1+d \end{bmatrix} = abcd\,(1+a^{-1}+b^{-1}+c^{-1}+d^{-1})$

f) $det\begin{bmatrix} x+a & a & c & d \\ a & x+b & c & d \\ a & b & x+c & d \\ a & b & c & x+d \end{bmatrix} = x^3(x+a+b+c+d)$

g) $det\begin{bmatrix} a^2 & a & 1 & bcd \\ b^2 & b & 1 & cda \\ c^2 & c & 1 & dab \\ d^2 & d & 1 & abc \end{bmatrix} = det\begin{bmatrix} a^3 & a^2 & a & 1 \\ b^3 & b^2 & b & 1 \\ c^3 & c^2 & c & 1 \\ d^3 & d^2 & d & 1 \end{bmatrix}$

23. Calcule los siguientes determinantes llevando la matriz correspondiente a su forma escalonada (y aplicando en cada paso los teoremas 2.4, 2.5, 2.6 y 2.7).

a) $det\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 4 & 3 & 2 \end{bmatrix}$

b) $det\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

c) $det\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$

d) $det\begin{bmatrix} 1 & 3 & 5 & 1 \\ -1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 0 & 0 & 2 & -1 \end{bmatrix}$

e) $det\begin{bmatrix} 0 & 4 & 10 & 0 \\ 0 & 0 & 4 & 9 \\ 4 & 6 & 8 & 8 \\ 0 & 0 & 0 & 12 \end{bmatrix}$

f) $det\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$

g) $det\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & -2 & 2 \\ 1 & 2 & -1 & 1 \\ 3 & 6 & -3 & 3 \end{bmatrix}$

h) $det\begin{bmatrix} 3 & 5 & 11 & 16 & 21 \\ 2 & 3 & 7 & 10 & 13 \\ 1 & 2 & 3 & 4 & 5 \\ 2 & -7 & 7 & 7 & 2 \\ 1 & 4 & 5 & 3 & 10 \end{bmatrix}$

3. EL DETERMINANTE DE UN PRODUCTO DE MATRICES

Esta sección será muy corta. Resulta que su contenido merece la pena de ser separado de resto del material que se está estudiando. El objetivo es establecer solamente una fórmula, una de las fórmulas más bellas del álgebra lineal.

Al estudiar la multiplicación de matrices y el cálculo de determinantes, puede surgir naturalmente la siguiente pregunta: ¿qué relación tiene el determinante de un producto de matrices cuadradas con los determinantes de cada una de las matrices en el producto?

Siendo tanto la multiplicación de matrices como el cálculo de determinantes operaciones altamente no triviales (a primera vista), en principio no se podría esperar una relación sencilla entre $det\,(AB)$ y los determinantes $det\,A$ y $det\,B$.

Por ejemplo, para la suma de las matrices cuadradas A y B (operación mucho más sencilla que la multiplicación de ellas) uno se puede plantear la misma pregunta anterior. Por supuesto que sería agradable que aconteciera $det\,(A + B) = det\,A + det\,B$. Sin embargo, esto es falso.

Considérese por ejemplo:

$$A = \begin{bmatrix} 5 & 8 \\ 2 & 10 \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 9 \\ 4 & 7 \end{bmatrix}$$

$$A + B = \begin{bmatrix} 4 & 17 \\ 6 & 17 \end{bmatrix}$$

Se tiene

$$det\,A = det \begin{bmatrix} 5 & 8 \\ 2 & 10 \end{bmatrix} = 34$$

$$det\,B = det \begin{bmatrix} -1 & 9 \\ 4 & 7 \end{bmatrix} = -43$$

$$det\,(A + B) = det \begin{bmatrix} 4 & 17 \\ 6 & 17 \end{bmatrix} = -34 \neq det\,A + det\,B$$

Ante esta perspectiva parecería ilusorio esperar una relación sencilla para el producto.

Considérense las mismas matrices del ejemplo anterior.

Se tiene

$$AB = \begin{bmatrix} 5 & 8 \\ 2 & 10 \end{bmatrix} \begin{bmatrix} -1 & 9 \\ 4 & 7 \end{bmatrix} = \begin{bmatrix} 27 & 101 \\ 38 & 88 \end{bmatrix}$$

Entonces

$$det\,(AB) = det \begin{bmatrix} 27 & 101 \\ 38 & 88 \end{bmatrix} = -1462 = (34)(-43) = (det\,A)(det\,B)$$

¿Coincidencia?

Afortunadamente no. Resultará que, en general, si se multiplican dos matrices cuadradas A y B y se calcula el determinante del producto AB, lo que dará por resultado el producto de los determinantes de A y de B, es decir,

$$det(AB) = (det\,A)(det\,B)$$

Es de esperarse que el lector pueda apreciar la sencillez y la elegancia de la fórmula anterior, para la cual se verá un argumento también sumamente sencillo, que establecerá su validez.

Primeramente se establecen un par de lemas preliminares.

LEMA 1

i) Si E es una matriz elemental que provino de I_n intercambiando en ésta dos de sus líneas, entonces $det\,E = -1$.
ii) Si E es una matriz elemental que provino de I_n multiplicando en ésta una de sus líneas por c, entonces $det\,E = c$.
iii) Si E es una matriz elemental que provino de I_n sustituyendo en ésta una de sus líneas por ella misma más un múltiplo de otra línea, entonces $det\,E = 1$.

DEMOSTRACIÓN i), ii) e iii) son consecuencia inmediata de los teoremas 2.4, 2.5 y 2.6, respectivamente, y del hecho de que $det\,I_n = 1$.

Q.E.D.

LEMA 2

Sea A una matriz de orden n y sea E una matriz elemental. Entonces $det(EA) = (detE)(detA)$.

DEMOSTRACIÓN Si A' es la matriz que se obtiene de A realizando en ésta una operación elemental, entonces, por el teorema 4.5 (capítulo 1), $A' = EA$, donde E es la matriz elemental que provino de I_n realizando en ésta la misma operación elemental. Entonces el lema se sigue inmediatamente de los teoremas 2.4, 2.5 y 2.6 junto con el lema (1).

Q.E.D.

TEOREMA 3.1

Sean A y B dos matrices cuadradas. Entonces

$$det(AB) = (detA)(detB)$$

DEMOSTRACIÓN Se van a considerar dos casos.

Caso 1. A es una matriz no inversible.

En este caso se afirma que AB también es una matriz no inversible, independientemente de cuál sea la matriz B. En efecto, supóngase por contradicción que AB es inversible. En tal caso existiría una matriz C tal que $(AB)C = I_n$. O sea, $A(BC) = I_n$. Pero esta última expresión nos dice que A es inversible (según el teorema 4.4 del capítulo 1) y que

$A^{-1} = BC$, una contradicción. Entonces, por el corolario del teorema 2.7, se tiene $det\,A = 0$, y $det(AB) = 0$. En este caso se cumple el resultado del teorema (ambos miembros de la igualdad son cero).

Caso 2. A es una matriz inversible.

En este caso, según el teorema 4.8 (capítulo 1), A puede escribirse como producto de matrices elementales, dígase

$$A = E_1 E_2 \dots E_k$$

Entonces

$$det\,(AB) = det\,(E_1E_2 \dots E_kB) = det\,(E_1E_2E_3 \dots E_kB)) \overset{\text{lema 2}}{=} (det\,E_1)\, det\,(E_2E_3 \dots E_kB)$$

$$\overset{\text{lema 2}}{=} (det\,E_1)(det(E_2(E_3E_4 \dots E_kB))) = (det\,E_1)(det\,E_2)\,det\,(E_3E_4 \dots E_kB)$$
$$= \dots = (det\,E_1)(det\,E_2) \dots (det\,E_k)(det\,B)$$

Pero por un argumento similar al anterior (aplicaciones sucesivas del lema 2)

$$det\,A = (det\,E_1)(det\,E_2) \dots (det\,E_k)$$

Entonces,

$$det\,(AB) = (det\,A)(det\,B)$$

Q.E.D.

Una consecuencia inmediata de este resultado es que se puede establecer (como ya se mencionaba al principio del capítulo) una caracterización de la propiedad de inversibilidad de una matriz en términos del valor de su determinante.

COROLARIO 1

La matriz A de orden n es inversible si, y sólo si $det\,A \neq 0$.

DEMOSTRACIÓN La parte "si" ya fue demostrada en el corolario del teorema 2.7. Se probará entonces el "sólo si". Supóngase que A es inversible. Entonces existe A^{-1} tal que $AA^{-1} = I$. Por tanto,

$$det\,(AA^{-1}) = det\,I$$
$$\parallel$$
$$(det\,A)(det\,A^{-1}) = 1$$

o sea, que $det\,A$ es un factor de un producto que da por resultado 1. Entonces $det\,A \neq 0$.

Q.E.D.

COROLARIO 2

Si A es una matriz inversible

$$(det\,A^{-1}) = (det\,A)^{-1} = \frac{1}{det\,A}$$

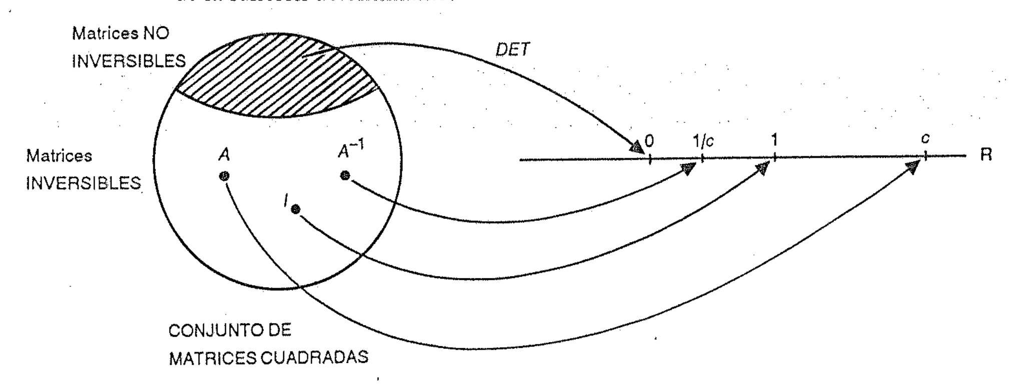
DEMOSTRACIÓN Consecuencia de la fórmula

$$(det\,A)(det\,A^{-1}) = 1$$

establecida en el corolario (1).

Q.E.D.

Los resultados anteriores permiten tener la siguiente representación esquemática de la función determinante:



EJERCICIOS (SECCIÓN 3, CAPÍTULO 2)

1. Considere las matrices

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 4 & 5 \\ -2 & 1 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & -1 & -1 \\ 1 & 2 & 1 \\ 4 & 2 & 6 \end{bmatrix}$$

a) Verifique que $det(AB) = (detA)(detB)$.

b) Sea $C = (c_{ij})_{i,j=1,2,3}$ la matriz cuyos elementos son:

$$c_{ij} = a_{i1}b_{j1} + a_{i2}a_{j2} + a_{i3}a_{j3} \qquad i,j = 1, 2, 3$$

Es decir, la matriz C se obtiene "multiplicando líneas de A por líneas de B". Por ejemplo, el elemento c_{23} se obtiene multiplicando la segunda línea de A por la tercera línea de B

$$c_{23} = (1)(4) + (4)(2) + (5)(6) = 42$$

Demuestre que $det\ C = (detA)(detB)$.

c) Sea $D = (d_{ij})_{i,j=1,2,3}$ la matriz cuyos elementos son:

$$d_{ij} = a_{1i}b_{1j} + a_{2i}a_{2j} + a_{3i}b_{3j} \qquad i, j = 1, 2, 3$$

Es decir, la matriz D se obtiene "multiplicando columnas de A por columnas de B". Por ejemplo, el elemento d_{31} se obtiene multiplicando la tercera columna de A por la primera columna de B

$$d_{31} = (3)(3) + (5)(1) + (2)(4) = 23$$

Compruebe que $det\ D = (detA)(detB)$.

① 2. (Generalización del ejercicio anterior.) Sean $A = (a_{ij})_{i,j=1,\ldots,n}$ y $B = (b_{ij})_{i,j=1,\ldots,n}$ dos matrices cualesquiera de orden n. Sea $C = (c_{ij})_{i,j=1,\ldots,n}$ la matriz cuyos elementos son:

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{jk} \qquad i, j = 1, 2, \ldots, n$$

y sea $D = (d_{ij})_{i,j=1,\ldots,n}$ la matriz cuyos elementos son:

$$d_{ij} = \sum_{k=1}^{n} a_{ki}b_{kj} \qquad i, j = 1, 2, \ldots, n$$

Demuestre que $det(AB) = (detA)(detB) = det\ C = det\ D$.

① 3. Sin calcular el determinante, pruebe que

$$det \begin{bmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{bmatrix} = (a^2 + b^2 + c^2 + d^2)^2$$

(*Sugerencia*: use el resultado del ejercicio anterior.)

4. Sea A una matriz de orden n. Demuestre que el determinante de la matriz A^2 es un número no negativo.

5. Al usar el teorema 3.1, dé una nueva demostración (véase ejercicio 14 de la sección anterior) de que

$$det\ (cA) = c^n\ det\ A$$

en donde A es una matriz de orden n y c es un número real cualquiera.

(*Sugerencia*: considere la matriz cI_n.)

6. Sea A una matriz de orden 3 cuyo determinante es igual a 2.

a) Compruebe que A es inversible.

b) Calcule $det\ A^{-1}$.

c) Calcule $det(5A^{-1})$.

d) Calcule $det(5A)^{-1}$.

e) Calcule $det(2A^{-1})^4$.

7. Sean A y B dos matrices de orden 4 tales que $det\ A = 3$ y $det\ B = -2$.

a) Demuestre que el producto AB es una matriz inversible. ¿Es el producto BA también inversible?

b) Calcule $det(AB)^{-1}$, y $det(BA)^{-1}$.

c) Calcule $det\ A^{-1}B$.

d) Calcule $det(2A^{-1})(3B^{-1})$.

e) Calcule $det(5B^{-1})(2A)(6B)$.

f) Calcule $det(5B)^{-1}(2A)(6B)$.

8. Sea A_σ la matriz asociada a la permutación σ del conjunto $\{1, 2, \ldots, n\}$ (véase ejercicio 11 de la sección 1). Compruebe que A_σ es una matriz inversible. Determine A_σ^{-1}.

9. Demuestre que si A es una matriz antisimétrica de orden n, con n impar, entonces A no puede ser inversible. (*Sugerencia*: véase ejercicio 16 (k) de la sección 2.)

④ 10. Considere las matrices

$$M_1 = [1],\ M_2 = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix},\ M_3 = \begin{bmatrix} 6 & 7 & 8 \\ 9 & 10 & 11 \\ 12 & 13 & 14 \end{bmatrix},\ M_4 = \begin{bmatrix} 15 & 16 & \ldots & \\ & & \ldots & \\ & & \ldots & 30 \end{bmatrix}, \ldots$$

a) Demuestre que la traza de la matriz M_n es

$$tr\ M_n = \frac{n}{6}(2n^3 + n + 3)$$

y que este valor coincide con la suma de los elementos de la diagonal secundaria de M_n.*

b) Compruebe que M_n es una matriz no inversible para $n \geq 3$.

④ 11. Sea A la matriz de orden n cuyos elementos son $a_{ij} = i + j$, $i, j = 1, 2, \ldots, n$.

a) Calcule la traza de A y demuestre que su valor coincide con la suma de los elementos de su diagonal secundaria.*

b) Pruebe que A no es inversible para $n \geq 3$.

12. Sea A una matriz nilpotente (véase ejercicio 14 de la sección 3, capítulo 1). Calcule $det\ A$.

13. Sea A una matriz involutoria (véase ejercicio 26 de la sección 3, capítulo 1). Calcule los valores posibles para $det\ A$.

4. DESARROLLO POR COFACTORES

Considere la matriz de orden 3

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

*La diagonal secundaria de la matriz $A = (a_{ij})_{i,j=1,\ldots,n}$ está constituida por los elementos a_{1n}, $a_{2,n-1}, \ldots, a_{n1}$.

Se ha visto que el determinante de esta matriz está dado por

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Al reescribir esta expresión factorizando los elementos de la primera línea de A se tiene

$$\det A = (a_{22}a_{33} - a_{23}a_{32})a_{11} + (a_{23}a_{31} - a_{21}a_{33})a_{12} + (a_{21}a_{32} - a_{22}a_{31})a_{13} \qquad (4.1)$$

Obsérvese que las expresiones entre paréntesis tienen un gran parecido al valor del determinante de una matriz de orden 2.

De hecho

$$\text{coeficiente de } a_{11} = a_{22}a_{33} - a_{23}a_{32} = \det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$$

$$\text{coeficiente de } a_{12} = a_{23}a_{31} - a_{21}a_{33} = -\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$$

$$\text{coeficiente de } a_{13} = a_{21}a_{32} - a_{22}a_{31} = \det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

Introdúzcase la notación

$$A(1, 1) = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$$

$$A(1, 2) = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$$

$$A(1, 3) = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

Véase entonces, que la fórmula (4.1) puede escribirse como

$$\det A = a_{11} \det A(1, 1) - a_{12} \det A(1, 2) + a_{13} \det A(1, 3) \qquad (4.2)$$

Más aún, obsérvese que $A(1, 1)$ —la matriz cuyo determinante es el coeficiente de a_{11} en $\det A$— es la submatriz que se obtiene de A "borrando" de ella la primera línea y la primera columna (que es precisamente la línea y la columna a la que pertenece el elemento a_{11}), o sea,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \rightarrow A(1, 1) = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$$

Similarmente

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \rightarrow A(1, 2) = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \rightarrow A(1, 3) = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

La fórmula (4.2) que establece el valor del determinante de A "con los elementos de la primera línea de A factorizados" se llama *desarrollo por cofactores del determinante de A respecto de su primera línea*. (Se verá que un resultado similar sigue siendo válido para cualquier línea o columna de A.) La ventaja más obvia que tiene esta fórmula es en relación a la facilidad de los cálculos para hallar el determinante de la matriz, pues permite evaluar el determinante de la matriz de orden 3 por medio de la evaluación de 3 determinantes de matrices de orden 2 —que se calculan muy fácilmente—.

EJEMPLO 1 Por ejemplo, si A es la matriz

$$A = \begin{bmatrix} 2 & 3 & 5 \\ -1 & 2 & -4 \\ 5 & 8 & 7 \end{bmatrix} \qquad (4.3)$$

se tiene que aplicando la fórmula (4.2)

$$\det A = 2 \det A(1, 1) - 3 \det A(1, 2) + 5 \det A(1, 3)$$
$$= 2 \det\begin{bmatrix} 2 & -4 \\ 8 & 7 \end{bmatrix} - 3 \det\begin{bmatrix} -1 & -4 \\ 5 & 7 \end{bmatrix} + 2 \det\begin{bmatrix} -1 & 2 \\ 5 & 8 \end{bmatrix}$$
$$= 2(46) - 3(13) + 5(-18) = -37$$

resultado que puede comprobarse haciendo el cálculo del determinante de A directamente con la definición.

El objetivo de esta sección es extender este tipo de ideas a matrices de orden arbitrario y obtener algunas consecuencias prácticas de los resultados que se establecerán.

Sea A una matriz de orden n con elementos a_{ij}.

Se denotará por $A(i, j)$ la submatriz de A, de orden $(n - 1)$, obtenida eliminando en A su i-ésima línea y su j-ésima columna. Es decir,

$$A(i,j) = \begin{bmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ & & \vdots & & & \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ & & \vdots & & & \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{bmatrix} \Bigg\} \begin{matrix} n-1 \\ \text{líneas} \end{matrix}$$

$n-1$ columnas

EJEMPLO 2 Por ejemplo, para la matriz (4.3) se tiene

$$A(2,2) = \begin{bmatrix} 2 & 5 \\ 5 & 7 \end{bmatrix} \quad A(3,1) = \begin{bmatrix} 3 & 5 \\ 2 & -4 \end{bmatrix} \quad A(3,3) = \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}$$

En el caso general (siendo A una matriz de orden n) se procederá análogamente a como se hizo al principio de la sección con la matriz de orden 3. La idea es expresar el determinante de A de tal manera que los elementos de su primera línea a_{11}, a_{12}, $\dots$, a_{1n}, aparezcan factorizados. Es decir, se procurará una expresión del tipo

$$\det A = \gamma_{11}a_{11} + \gamma_{12}a_{12} + \dots + \gamma_{1n}a_{1n} = \sum_{k=1}^{n} \gamma_{1k}a_{1k} \qquad (4.4)$$

Al coeficiente γ_{1i} que aparece multiplicando al elemento a_{1i} $(i=1, 2, \dots, n)$ en la expresión anterior se le llama *cofactor* (del elemento a_{1i}).

Se tiene, según la definición de determinante, que

$$\det A = \sum_{\pi} (sgn\ \pi) a_{1\,\pi(1)} a_{2\,\pi(2)} \dots a_{n\,\pi(n)}$$

Se llamará S_n al conjunto de todas las $n!$ permutaciones de $\{1, 2, \dots, n\}$. Se escribirá más explícitamente la fórmula anterior como

$$\det A = \sum_{\pi \in S_n} (sgn\ \pi) a_{1\,\pi(1)} a_{2\,\pi(2)} \dots a_{n\,\pi(n)} \qquad (4.5)$$

De los $n!$ sumandos que aparecen en la fórmula anterior, se agrupan aquellos que tienen al elemento a_{11} como factor común. Éstos son:

$$\sum_{\substack{\pi \in S_n \\ \pi(1)=1}} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & \pi(2) & \dots & \pi(n) \end{pmatrix} a_{11} a_{2\,\pi(2)} \dots a_{n\,\pi(n)}$$

Similarmente, la fórmula

$$\sum_{\substack{\pi \in S_n \\ \pi(1)=2}} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ 2 & \pi(2) & \dots & \pi(n) \end{pmatrix} a_{12} a_{2\,\pi(2)} \dots a_{n\,\pi(n)}$$

exhibe todos los sumandos de (4.5) en los que aparece a_{12} como factor común. En general, la parte de la fórmula (4.5) en la que se exhibe explícitamente el factor a_{1k} es:

$$\sum_{\substack{\pi \in S_n \\ \pi(1)=k}} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ k & \pi(2) & \dots & \pi(n) \end{pmatrix} a_{1k} a_{2\,\pi(2)} \dots a_{n\,\pi(n)}$$

Entonces, la fórmula (4.5) puede escribirse como

$$\begin{aligned} \det A = & \left(\sum_{\substack{\pi \in S_n \\ \pi(1)=1}} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & \pi(2) & \dots & \pi(n) \end{pmatrix} a_{2\,\pi(2)} \dots a_{n\,\pi(n)} \right) a_{11} + \\ & + \dots + \left(\sum_{\substack{\pi \in S_n \\ \pi(1)=k}} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ k & \pi(2) & \dots & \pi(n) \end{pmatrix} a_{2\,\pi(2)} \dots a_{n\,\pi(n)} \right) a_{1k} + \\ & + \dots + \left(\sum_{\substack{\pi \in S_n \\ \pi(1)=n}} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ n & \pi(2) & \dots & \pi(n) \end{pmatrix} a_{2\,\pi(2)} \dots a_{n\,\pi(n)} \right) a_{1n} \\ = & \sum_{k=1}^{n} \left[\sum_{\substack{\pi \in S_n \\ \pi(1)=k}} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ k & \pi(2) & \dots & \pi(n) \end{pmatrix} a_{2\,\pi(2)} \dots a_{n\,\pi(n)} \right] a_{1k} \end{aligned}$$

Al comparar esta última fórmula con (4.4) se ve que

$$\gamma_{1k} = \sum_{\substack{\pi \in S_n \\ \pi(1)=k}} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ k & \pi(2) & \dots & \pi(n) \end{pmatrix} a_{2\,\pi(2)} \dots a_{n\,\pi(n)} \qquad (4.6)$$

Se quiere ver que en realidad γ_{1k} es (excepto quizás por un signo) el valor del determinante de la matriz $A(1, k)$. Más concretamente, se afirma que

$$\gamma_{1k} = (-1)^{1+k} \det A(1, k) \qquad (4.7)$$

Se probará, primeramente, que esta fórmula es cierta para $k = 1$; es decir, que

$$\gamma_{11} = \det A(1, 1)$$

En efecto, según la fórmula (4.6) se tiene que

$$\gamma_{11} = \sum_{\substack{\pi \in S_n \\ \pi(1)=1}} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & \pi(2) & \dots & \pi(n) \end{pmatrix} a_{2\,\pi(2)} \dots a_{n\,\pi(n)}$$

Primeramente obsérvese que en el término $a_{2\pi(2)} \dots a_{n\pi(n)}$ de esta última fórmula, no aparece ningún elemento de la primera línea de A (el primer subíndice es diferente de 1 en todos los factores) así como tampoco aparece ningún elemento de su primera columna (el segundo subíndice es $\pi(j)$ para $\pi \in S_n$; pero $\pi(1) = 1$ y entonces $\pi(j) \neq 1$ para $j = 2, \dots, n$). Éstos son, pues, productos de $n - 1$ elementos de $A(1, 1)$ cada uno de los cuales proviene de una línea y de una columna diferentes (son entonces productos elementales). Véase, además, que para cada permutación $\pi \in S_n$ tal que $\pi(1) = 1$, le corresponde de manera obvia una permutación $\pi' \in S_{n-1}$ del conjunto $\{2, 3, \dots, n\}$ (defina $\pi'(j) = \pi(j), j = 2, 3, \dots n$). Es claro también que dada una permutación $\pi' \in S_{n-1}$ del conjunto $\{2, 3, \dots, n\}$, puede asociársele a ésta la permutación $\pi \in S_n$ tal que $\pi(1) = 1$ y $\pi(j) = \pi'(j)\ j = 2, \dots, n$.

Existe, pues, una correspondencia uno a uno entre las permutaciones $\pi \in S_n$ tales que $\pi(1) = 1$ y las permutaciones $\pi' \in S_{n-1}$ del conjunto $\{2, 3, \dots, n\}$. Además, la permutación $\pi' \in S_{n-1}$ de $\{2, \dots, n\}$ tendrá una inversión si, y sólo si la permutación correspondiente $\pi \in S_n$, $\pi(1) = 1$ la tiene (pues ninguna de las inversiones de $(1, \pi(2), \dots, \pi(n))$ involucra al elemento 1). Por lo tanto,

$$det\, A(1, 1) = \sum_{\pi \in S_{n-1}} sgn \begin{pmatrix} 2 & \dots & n \\ \pi'(2) & \dots & \pi'(n) \end{pmatrix} a_{2\,\pi'(2)} \dots a_{n\,\pi'(n)}$$

$$= \sum_{\substack{\pi \in S_n \\ \pi(1)=1}} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & \pi(2) & \dots & \pi(n) \end{pmatrix} a_{2\,\pi(2)} \dots a_{n\,\pi(n)} = \gamma_{11}$$

y queda así establecida la fórmula (4.7) para $k = 1$. Hasta este momento se tiene que

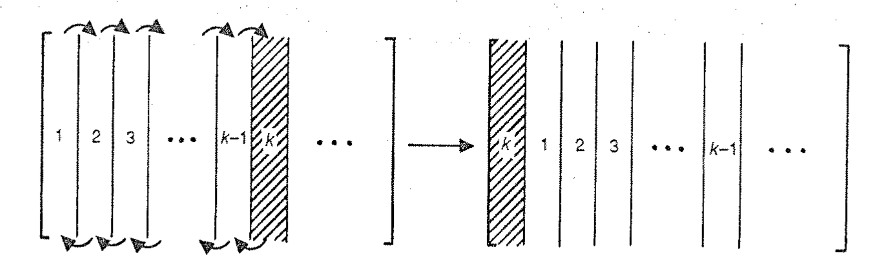
$$det\, A = (det\, A(1, 1))a_{11} + \gamma_{12}a_{12} + \dots + \gamma_{1n}a_{1n} \tag{4.8}$$

Antes de proceder a probar la fórmula (4.7) para un k arbitrario, se desea dejar establecido "con palabras" lo que se ha ya probado:

> Sea A una matriz de orden n. El factor que multiplica al elemento de la primera línea y primera columna, en el desarrollo del determinante de A, es el valor del determinante de la submatriz de orden $(n - 1)$ que se obtiene de A eliminando su primera línea y su primera columna.
>
> (4.9)

Sea ahora $k > 1$. Procédase a intercambiar la posición de la k-ésima columna de A con la posición de su primera columna, *haciendo sólo intercambio entre columnas adyacentes.*

Esquemáticamente:



Obsérvese que se ha logrado dejar en el mismo orden relativo el resto de las columnas de A. Llámese $A' = (a'_{ij})_{i,j=1,\dots,n}$ a la matriz obtenida por este intercambio de columnas. Según lo visto en la subsección 2.3, se tiene que

$$det\, A' = (-1)^{k-1}\, det\, A$$

o bien,

$$det\, A = (-1)^{k+1}\, det\, A'$$

Por otra parte, según lo "hablado" en (4.9) se tiene que

$$det\, A' = (det\, A'(1, 1))a'_{11} + \dots$$

o sea,

$$det\, A' = (-1)^{1+k}\, det\, A' = ((-1)^{1+k}\, det\, A'(1, 1)a'_{11} + \dots$$

Pero $A'(1, 1) = A(1, k)$ y $a'_{11} = a_{1k}$. Entonces,

$$det\, A = ((-1)^{1+k}\, det\, A(1, k))a_{1k} + \dots$$

Al comparar esta expresión con la fórmula (4.8) (igualando los coeficientes de a_{1k} en ambas expresiones) se obtiene

$$\gamma_{1k} = (-1)^{1+k}\, det\, A(1, k)$$

que era lo que se quería demostrar.

En conclusión, se ha demostrado el siguiente teorema:

TEOREMA 4.1

> (DESARROLLO POR COFACTORES RESPECTO DE LA PRIMERA LÍNEA DE UNA MATRIZ.)
> Sea A una matriz de orden n de elementos a_{ij}. Sea $A(1, k)$ la submatriz de orden $(n - 1)$ que se obtiene de A eliminando de ésta su primera línea y su k-ésima columna. Entonces
>
> $$det\, A = (det\, A(1, 1))a_{11} + \dots + ((-1)^{1+k}\, det\, A(1, k)))a_{1k} + \dots + ((-1)^{1+n}\, det\, A(1, n))a_{1n} \tag{4.10}$$

(el número $\gamma_{1k} = (-1)^{1+k} \det A(1, k)$, $k = 1, 2, \ldots, n$ es el *cofactor*, correspondiente al elemento a_{1k}).

En realidad, este teorema es un caso particular del siguiente resultado más general.

TEOREMA 4.2

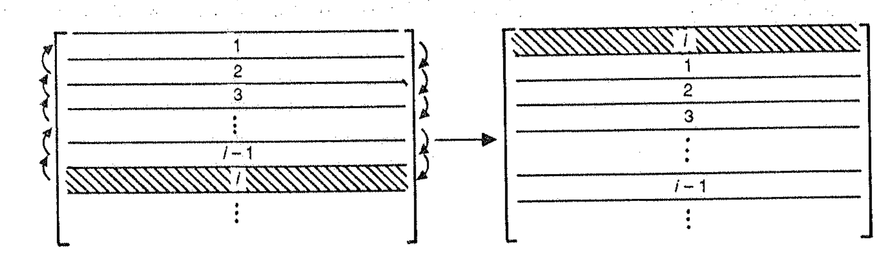
(DESARROLLO POR COFACTORES RESPECTO DE LA *I*-ÉSIMA LÍNEA DE UNA MATRIZ.)
Sea A una matriz de orden n de elementos a_{ij}. Sea $A(i, k)$ la submatriz de orden $(n - 1)$ que se obtiene de A eliminando en ésta su i-ésima línea y su k-ésima columna. Entonces

$$\det A = ((-1)^{i+1} \det A(i, 1))a_{i1} + \ldots + ((-1)^{i+n} \det A(i, n))a_{in}$$

$$= \sum_{k=1}^{n} ((-1)^{i+k} \det (A(i, k))a_{ik}$$

(el número $\gamma_{ik} = (-1)^{i+k} \det A(i, k)$ es el cofactor asociado al elemento a_{ik}).

DEMOSTRACIÓN Intercámbiese la i-ésima línea de A con su primera línea, realizando sólo intercambios de líneas adyacentes (ésto tomará entonces $(i - 1)$ intercambios), como se muestra esquemáticamente enseguida:



Sea $A' = (a'_{ij})_{i,j=1,\ldots,n}$ la matriz obtenida después de este cambio de líneas de A. Por una parte se tiene

$$\det A' = (-1)^{i-1} \det A$$

o sea,

$$\det A = (-1)^{i+1} \det A'$$

Por otra parte, se puede aplicar el teorema 4.1 a la matriz A' obteniendo

$$\det A' = \sum_{k=1}^{n} ((-1)^{1+k} \det A'(1, k))a'_{1k}$$

o sea que

$$\det A = (-1)^{i+1} \det A' = \sum_{k=1}^{n} ((-1)^{1+k} \det A'(1, k))a'_{1k}$$

Pero $A'(1, k) = A(i, k)$ y $a'_{1k} = a_{ik}$. Entonces

$$\det A = \sum_{k=1}^{n} ((-1)^{i+k} \det A(1, k))a_{ik}$$

como se quería.

Q.E.D.

El desarrollo por cofactores del determinante de una matriz A, puede hacerse también respecto de las columnas de ésta. El siguiente teorema hace explícito este hecho. Su demostración debe ser, usando el teorema 4.2, un ejercicio elemental que se deja para el lector.

TEOREMA 4.3

(DESARROLLO POR COFACTORES RESPECTO DE LA *J*-ÉSIMA COLUMNA DE UNA MATRIZ.)
Se tiene la siguiente columna para $\det A$:

$$\det A = ((-1)^{1+j} \det A(1, j))a_{1j} + \ldots + ((-1)^{n+j} \det A(n, j))a_{nj}$$

$$= \sum_{k=1}^{n} ((-1)^{k+j} \det A(k, j))a_{kj}$$

Considérese por ejemplo, la misma matriz (4.3) y realícese el cálculo de su determinante respecto de la 3a. columna. Se obtiene

$$\det A = \sum_{k=1}^{3} ((-1)^{k+3} \det A(k, 3))a_{k3}$$

$$= (-1)^{1+3} \det A(1, 3)a_{13} + (-1)^{2+3} \det A(2, 3)a_{23} + (-1)^{3+3} \det A(3, 3)a_{33}$$

$$= \left(\det\begin{bmatrix} -1 & 2 \\ 5 & 8 \end{bmatrix}\right)(5) - \left(\det\begin{bmatrix} 2 & 3 \\ 5 & 8 \end{bmatrix}\right)(-4) + \left(\det\begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}\right)(7)$$

$$= (-18)(5) - (1)(-4) + (7)(7) = -37$$

Al determinante $\det A(i, j)$ se le llama *menor* de A, de orden $n - 1$. Con esta terminología, se ha visto que el determinante de A puede expresarse como una suma de los elementos de su i-ésima línea o j-ésima columna multiplicados por $(-1)^{i+j}$ veces sus correspondientes menores de orden $n - 1$.

Una manera fácil de recordar los signos que tienen estos menores (para dar así los cofactores correspondientes) es observando que éstos están distribuidos alternadamente en la matriz A: el menor de a_{11} tiene signo $(-1)^{1+1} = +1$, el de a_{12} es $(-1)^{1+2} = -1$, etc., quedando entonces distribuidos como sigue:

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ & & --- & & \end{bmatrix}$$

Desde el punto de vista computacional, el desarrollo por cofactores de un determinante de una matriz de orden n no tienen en principio ventaja alguna respecto del cálculo del determinante usando directamente la definición de él: el cálculo del determinante de la matriz se reduce al cálculo de n determinantes de matrices de orden $(n - 1)$; por tanto, habría que realizar $(n)(n - 1)! = n!$ operaciones para tal cálculo, el mismo número de operaciones que si se evalúa el determinante usando directamente su definición. Sin embargo, se pueden combinar las ideas expuestas en la subsección 2.3 con el desarrollo por cofactores, procediendo primeramente a hacer operaciones con las líneas y/o con las columnas de la matriz, de tal forma que se logre la mayor cantidad de ceros posible en alguna de las líneas o columnas, y luego desarrollando por cofactores con respecto a esa línea o columna. De esta forma se puede ahorrar el cálculo de los menores correspondientes a los elementos que son cero (pues en la expresión del desarrollo por cofactores aparecen éstos multiplicados por el menor correspondiente).

Véase un ejemplo.

EJEMPLO 3

Suponga que se quiere calcular el determinante

$$D = det\begin{bmatrix} 2 & 1 & 0 & 1 & 8 & 4 & 0 \\ 1 & 4 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 8 \\ 3 & 0 & 1 & 0 & 0 & 0 & 9 \\ 4 & 2 & 5 & 1 & 2 & 0 & 7 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Al aplicar la definición de determinante, se tendría que realizar $7! = 5040$ operaciones para calcular D. Sin embargo, se verá que procediendo inteligentemente se puede ahorrar mucho trabajo.

En principio, vea que la última línea de la matriz tiene todos sus elementos iguales a cero excepto el último de ellos (que es 1). Conviene, entonces, hacer el desarrollo de D por cofactores respecto de esa línea, pues así se reduce el cálculo a

$$D = (1)(-1)^{7+7} det\begin{bmatrix} 2 & 1 & 0 & 1 & 8 & 4 \\ 1 & 4 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 & 0 \\ 4 & 2 & 5 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} = det\begin{bmatrix} 2 & 1 & 0 & 1 & 8 & 4 \\ 1 & 4 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 & 0 \\ 4 & 2 & 5 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

a_{77} cofactor de $a_{77} = (-1)^{7+7} det\, A(7, 7)$

La columna con mayor número de ceros es la 6a. columna. Si desarrolla D por cofactores respecto de esa columna obtendrá sólo 2 (de los 6) sumandos diferentes de cero. Sin embargo, si sustituye la primera línea por ella misma menos 4 veces la sexta línea, el determinante no altera su valor y quedaría

$$D = det\begin{bmatrix} 2 & -3 & -4 & 1 & 8 & 0 \\ 1 & 4 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 & 0 \\ 4 & 2 & 5 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

de tal modo que ahora la última columna tiene sólo un elemento distinto de cero. Al desarrollar por cofactores respecto de esa columna se obtiene

$$D = (1)(-1)^{6+6} det\begin{bmatrix} 2 & -3 & -4 & 1 & 8 \\ 1 & 4 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 \\ 4 & 2 & 5 & 1 & 2 \end{bmatrix} = det\begin{bmatrix} 2 & -3 & -4 & 1 & 8 \\ 1 & 4 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 \\ 4 & 2 & 5 & 1 & 2 \end{bmatrix}$$

Nuevamente, si sustituye la primera línea por ella misma menos la quinta línea, el valor D no se altera y se obtiene

$$D = det\begin{bmatrix} -2 & -5 & -9 & 0 & 6 \\ 1 & 4 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 \\ 4 & 2 & 5 & 1 & 2 \end{bmatrix}$$

Ahora la cuarta columna tiene sólo un elemento no nulo. Desarrolle por cofactores respecto de esa columna para obtener

$$D = (1)(-1)^{5+4} det\begin{bmatrix} -2 & -5 & -9 & 6 \\ 1 & 4 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 0 & 1 & 0 \end{bmatrix} = -\, det\begin{bmatrix} -2 & -5 & -9 & 6 \\ 1 & 4 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 0 & 1 & 0 \end{bmatrix}$$

Sustituya ahora la primera línea por ella menos 6 veces la segunda y obtenga

$$D = -\det\begin{bmatrix} -8 & -29 & -15 & 0 \\ 1 & 4 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 0 & 1 & 0 \end{bmatrix}$$

Al desarrollar D por cofactores respecto de la última columna queda

$$D = -(1)(-1)^{2+4}\det\begin{bmatrix} -8 & -29 & -15 \\ 1 & 1 & 1 \\ 3 & 0 & 1 \end{bmatrix} = -\det\begin{bmatrix} -8 & -29 & -15 \\ 1 & 1 & 1 \\ 3 & 0 & 1 \end{bmatrix}$$

Sustituya la primera columna por ella misma menos 3 veces la tercera y obtenga

$$D = -\det\begin{bmatrix} 37 & -29 & -15 \\ -2 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Al desarrollar por cofactores respecto de la tercera línea queda finalmente

$$D = -(1)(-1)^{3+3}\det\begin{bmatrix} 37 & -29 \\ -2 & 1 \end{bmatrix} = -\det\begin{bmatrix} 37 & -29 \\ -2 & 1 \end{bmatrix} = -((37)(1)-(-2)(-29)) = 21$$

Los resultados obtenidos en esta sección tienen también importancia desde el punto de vista teórico. En la próxima sección se usarán y se obtendrán varios resultados nuevos interesantes.

EJERCICIOS (SECCIÓN 4, CAPÍTULO 2)

1. Sea A la matriz

$$A = \begin{bmatrix} 1 & 3 & 2 \\ -1 & 4 & 5 \\ 2 & 1 & 8 \end{bmatrix}$$

Sea C_{ij} el cofactor asociado al elemento a_{ij} de la matriz A. Calcule C_{12}, C_{31}, C_{22}, C_{33} y C_{13}.

2. Calcule el determinante

$$\det\begin{bmatrix} 5 & -1 & 2 \\ 3 & 2 & 4 \\ -1 & 2 & -2 \end{bmatrix}$$

desarrollando por cofactores respecto de

a) la primera línea
b) la segunda columna
c) la tercera línea
d) la primera columna

3. Considere el determinante

$$\det\begin{bmatrix} 1 & 0 & 4 & -1 \\ 3 & 0 & 2 & 1 \\ x & a & y & z \\ 3 & 0 & 2 & 8 \end{bmatrix}$$

a) ¿Cuánto vale este determinante si $a = 0$?
b) ¿Cuánto vale si $a \neq 0$?
c) ¿Por qué el valor de este determinante no depende de x, y, z?

4. Al usar la fórmula del desarrollo por cofactores de un determinante, demuestre que el determinante de una matriz con una línea o con una columna de ceros vale cero (teorema 2.7).

5. Calcule los siguientes determinantes:

a) $\det\begin{bmatrix} 2 & 1 & 3 \\ 0 & 4 & 0 \\ 1 & 8 & 2 \end{bmatrix}$

b) $\det\begin{bmatrix} 3 & 1 & 5 \\ 2 & 1 & 4 \\ 0 & 4 & 2 \end{bmatrix}$

c) $\det\begin{bmatrix} -1 & 3 & -1 & 2 \\ 1 & -10 & 3 & 1 \\ 0 & 0 & 3 & 0 \\ 6 & 7 & -1 & 2 \end{bmatrix}$

d) $\det\begin{bmatrix} -1 & 3 & 8 & 2 \\ 1 & -10 & 18 & 1 \\ 0 & 0 & 3 & 0 \\ 6 & 7 & 46 & 2 \end{bmatrix}$ (compare con c))

e) $\det\begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 3 & 8 & 2 & 1 & 7 \\ 4 & 47 & 29 & 1 & 9 \\ -2 & 170 & 80 & -1 & -5 \\ 1 & 90 & 100 & 0 & 2 \end{bmatrix}$

f) $\det\begin{bmatrix} 2 & 0 & 1 & 0 & 4 \\ 4 & 1 & 2 & 3 & 7 \\ 0 & 0 & 1 & 0 & 3 \\ 2 & 1 & 3 & 1 & 9 \\ 3 & 1 & -1 & 1 & 3 \end{bmatrix}$

g) $\det\begin{bmatrix} 5 & 2 & 0 & 1 & 2 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 4 & 3 & 1 & 2 & 7 & 1 \\ 9 & 6 & 2 & 4 & 15 & 3 \\ 1 & 1 & 0 & 2 & 3 & 1 \\ 13 & 9 & 3 & 6 & 21 & 5 \end{bmatrix}$

③ 6. Demuestre la validez de las siguientes expresiones (las matrices involucradas son de orden n)

a) $\det\begin{bmatrix} 2 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 2 & 1 & \cdots & 0 & 0 \\ & & & \ddots & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 & 2 \end{bmatrix} = n + 1$

b) $det\begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ -1 & 1 & 1 & \dots & 1 & 1 \\ 0 & -1 & 1 & \dots & 1 & 1 \\ & & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix} = 2^{n-1}$

c) $det\begin{bmatrix} 1+x^2 & x & 0 & \dots & 0 & 0 \\ x & 1+x^2 & x & \dots & 0 & 0 \\ 0 & x & 1+x^2 & \dots & 0 & 0 \\ & & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & \dots & x & 1+x^2 \end{bmatrix} = 1 + x^2 + \dots + x^{2n}$

(*Sugerencia*: use inducción.)

7. Sea A la matriz

$$A = \begin{bmatrix} 2 & 4 & 0 & 0 \\ 1 & 6 & 0 & 0 \\ 0 & 0 & 5 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix}$$

a) Calcule *det A*.

b) Calcule el determinante de la matriz

$$A_1 = \begin{bmatrix} 2 & 4 & 0 & 0 \\ 1 & 6 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

y compare su valor con el determinante de la matriz

$$P = \begin{bmatrix} 2 & 4 \\ 1 & 6 \end{bmatrix}$$

c) Calcule el determinante de la matriz

$$A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix}$$

y compare su valor con el determinante de la matriz

$$Q = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$$

d) Verifique que $A = A_1A_2$.

e) Calcule $det\,(PQ)$ y verifique que $det\,A = det\,(PQ) = (det\,P)(det\,Q)$.

③ 8. (Generalización del ejercicio anterior.) Sean P y Q dos matrices cuadradas de orden n y considere la matriz A de orden $2n$

$$P = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$$

(en donde 0 es la matriz cero de orden n). Sean A_1 y A_2 las matrices (de orden $2n$) siguientes:

$$A_1 = \begin{bmatrix} P & 0 \\ 0 & I_n \end{bmatrix} \qquad A_2 = \begin{bmatrix} I_n & 0 \\ 0 & Q \end{bmatrix}$$

Compruebe que:

a) $A = A_1A_2$.

b) $det\,A_1 = det\,P$, $det\,A_2 = det\,Q$. Concluya entonces que $det\,A = (det\,P)(det\,Q)$.

9. Use el ejercicio anterior para calcular el determinante de la matriz

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 & 0 & 0 \\ 3 & 4 & 2 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

③ 10. Sean P, Q y R matrices cuadradas de orden n. Considere la matriz A de orden $2n$

$$A = \begin{bmatrix} P & R \\ 0 & Q \end{bmatrix}$$

Demuestre que $det\,A = (det\,P)(det\,Q)$.

④ 11. Sean P, Q, R y S matrices de orden n. Suponga que Q es una matriz inversible y que S conmuta con Q. Pruebe que

$$det\begin{bmatrix} P & R \\ S & Q \end{bmatrix} = det\,(PQ - RS)$$

(Observe que el miembro izquierdo de esta igualdad se refiere al determinante de una matriz de orden $2n$, mientras que el miembro derecho de la misma se refiere a un determinante de una matriz de orden n.)

12. Use el ejercicio anterior para calcular el determinante de la matriz

$$A = \begin{bmatrix} 1 & 3 & -1 & 3 & -1 & 2 \\ 0 & 2 & 5 & 2 & 1 & 1 \\ 3 & 0 & 0 & 4 & 2 & 0 \\ 3 & 1 & 0 & 1 & 2 & -1 \\ 0 & 3 & 1 & 0 & 1 & 2 \\ 0 & 0 & 3 & 0 & 0 & 1 \end{bmatrix}$$

⑤ 13. Demuestre el resultado del ejercicio 11 eliminando la hipótesis de la inversibilidad de la matriz Q.

14. Use el ejercicio anterior para demostrar que el determinante

$$det \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & a & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 2 & 2 & 2 & 0 & b & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 3 & 3 & 3 & 0 & 0 & c & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 4 & 4 & 0 & 0 & 0 & d & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 5 & 0 & 0 & 0 & 0 & e & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & f \\ 1 & 2 & 3 & 4 & 5 & 6 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 & 5 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 3 & 4 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

vale $a\,b\,c\,d\,e\,f$.

5. LA ADJUNTA DE UNA MATRIZ

El objetivo de esta sección es establecer, por una parte, una fórmula para la inversa de una matriz A y, por otra, obtener la conocida regla de Cramer para resolver sistemas de ecuaciones. Para lograr esto, se tendrá que introducir el concepto de "adjunta" de una matriz.

Primeramente se harán algunas observaciones concretas sobre la matriz

$$A = \begin{bmatrix} 2 & 5 & 1 \\ 7 & -1 & 3 \\ 2 & 4 & 1 \end{bmatrix} \qquad (5.1)$$

Se llamará C_{ij} al cofactor asociado al elemento a_{ij}.

Se tiene

$$C_{11} = (-1)^{1+1}\, det\, A(1,1) = det \begin{bmatrix} -1 & 3 \\ 4 & 1 \end{bmatrix} = -13$$

$$C_{12} = (-1)^{1+2}\, det\, A(1,2) = -det \begin{bmatrix} 7 & 3 \\ 2 & 1 \end{bmatrix} = -1$$

$$C_{13} = (-1)^{1+3}\, det\, A(1,3) = det \begin{bmatrix} 7 & -1 \\ 2 & 4 \end{bmatrix} = 30$$

Similarmente se obtienen

$C_{21} = -1$ $C_{31} = 16$

$C_{22} = 0$ $C_{32} = 1$

$C_{23} = 2$ $C_{33} = -37$

Se sabe que

desarrollo por cofactores respecto de las líneas de A:
$$\left\{\begin{array}{l} 2C_{11} + 5C_{12} + C_{13} \\ 7C_{21} - C_{22} + 3C_{23} \\ 2C_{31} + 4C_{32} + C_{33} \end{array}\right.$$

desarrollo por cofactores respecto de las columnas de A:
$$\left\{\begin{array}{l} 2C_{11} + 7C_{21} + 2C_{31} \\ 5C_{12} - C_{22} + 4C_{32} \\ C_{13} + 3C_{23} + C_{33} \end{array}\right.$$

$$= det\, A = -1$$

Uno se podría preguntar qué sucedería si se hace el desarrollo por cofactores con respecto de la línea i $(1 \le i \le 3)$ "desordenadamente", en el sentido de que cada sumando en el desarrollo no sea el producto del elemento a_{ij} por *su* cofactor asociado C_{ij}, sino que sea el producto de a_{ij} por el cofactor C_{kj} con $k \neq i$. Por ejemplo, si $i = 1$ y $k = 3$ se obtendría

$$a_{11}C_{31} + a_{12}C_{32} + a_{13}C_{33} = (2)(16) + (5)(1) + (1)(-37) = 0$$

Se podrían intentar algunas otras variantes y se vería que el resultado siempre será cero. Lo mismo sucedería si se hace este cálculo "desordenado" del desarrollo por cofactores respecto de alguna de sus columnas.

La razón de este hecho es muy simple y está explicada en la demostración del siguiente teorema:

TEOREMA 5.1

> Sea $A = (a_{ij})_{i,j=1,\ldots,n}$ una matriz cualquiera de orden n. Sea C_{ij} el cofactor asociado al elemento a_{ij}. Entonces
>
> $$\sum_{k=1}^{n} a_{rk}C_{sk} = \begin{cases} det\, A & \text{si } s = r \\ 0 & \text{si } s \neq r \end{cases}$$

DEMOSTRACIÓN El caso $s = r$ es el teorema 4.2 (desarrollo por cofactores con respecto de la r-ésima línea). Se demuestra entonces el caso $s \neq r$. Sea $A' = (a'_{ij})_{i,j=1,\ldots,n}$ la matriz de elementos $a'_{ij} = a_{ij}$ si $i \neq s$ y $a'_{sj} = a_{rj}$ $(j = 1, 2, \ldots, n)$. A' es pues una matriz que coincide con A en todos sus elementos excepto los de la s-ésima línea, en la cual A' tiene una copia de la r-ésima línea de A (por tanto, A' tiene dos líneas iguales, la r-ésima y la s-ésima). Se llamará C'_{ik} a los correspondientes cofactores de la matriz A'.

Calcúlese $det\, A'$ desarrollando por cofactores con respecto de su s-ésima línea. Se tiene

$$det\, A' = \sum_{k=1}^{n} a'_{sk}C'_{sk}$$

Obsérvese que $A'(s,k) = A(s,k)$ y por tanto $C_{sk} = C'_{sk}$. Además, $a'_{sk} = a_{rk}$, de modo que

$$\det A' = \sum_{k=1}^{n} a_{rk}C_{sk}$$

Pero (corolario del teorema 2.4) $\det A' = 0$.

Q.E.D.

Este teorema jugará un papel muy importante en la próxima subsección, en la cual se establecerá una fórmula para la inversa de una matriz.

Dada la matriz $A = (a_{ij})_{i,j=1,\ldots,n}$ se puede construir la *matriz de cofactores* de los elementos de A, poniendo $C = (C_{ij})_{i,j=1,\ldots,n}$ donde C_{ij} es el cofactor asociado al elemento a_{ij}.

EJEMPLO 1 Por ejemplo, para la matriz (5.1), su matriz C de cofactores es:

$$C = \begin{bmatrix} -13 & -1 & 30 \\ -1 & 0 & 2 \\ 16 & 1 & -37 \end{bmatrix} \tag{5.2}$$

DEFINICIÓN 5.1 Sea A una matriz de orden n. Se llama *matriz adjunta* de A, denotada por $adj\, A$, a la transpuesta de la matriz de cofactores de A.

La matriz $adj\, A$ es pues una matriz del mismo orden que A, que se obtiene intercambiando líneas por columnas en la matriz de cofactores de A.

EJEMPLO 2 Por ejemplo la matriz adjunta de (5.1) es la transpuesta de (5.2). Es decir,

$$adj\ A = \begin{bmatrix} -13 & -1 & 16 \\ -1 & 0 & 1 \\ 30 & 2 & -37 \end{bmatrix}$$

En general se tiene entonces que $adj\, A = (C_{ji})_{i,j=1,\ldots,n}$ donde C_{ij} es el cofactor asociado al elemento a_{ij} de la matriz A.

5.1. UNA FÓRMULA PARA LA INVERSA

El resultado principal de esta subsección se encuentra en el siguiente teorema:

TEOREMA 5.2 Sea A una matriz de orden n. Sea $adj\, A$ su matriz adjunta. Entonces

$$(A)\,(adj\, A) = (\det A)\, I$$

DEMOSTRACIÓN Escríbase $A = (a_{ij})_{i,j=1,\ldots,n}$
$adj A = (b_{ij})_{i,j=1,\ldots,n}$

(siendo así, $b_{ij} = C_{ji}$).

El elemento de la r-ésima línea y s-ésima columna de $A(adj A)$ es:

$$\sum_{k=1}^{n} a_{rk}b_{ks}$$

o sea,

$$\sum_{k=1}^{n} a_{rk}C_{sk}$$

Por el teorema 5.1, este elemento es diferente de cero (y vale $\det A$) sólo cuando $r = s$. Por tanto, el producto $A(adj\, A)$ se ve como

$$A\,(adj\, A) = \begin{bmatrix} \det A & & & 0 \\ & \det A & & \\ & & \ddots & \\ 0 & & & \det A \end{bmatrix} = (\det A)I_n$$

Q.E.D.

COROLARIO (UNA FÓRMULA PARA LA INVERSA.)
Si A es inversible, entonces

$$A^{-1} = \frac{1}{\det A}\, adj\, A \tag{5.3}$$

DEMOSTRACIÓN Por el teorema anterior se tiene

$$A(adj\, A) = (\det A)I_n$$

Si A es inversible se tiene $\det A \neq 0$ (corolario del teorema 3.1) y entonces se puede escribir

$$\frac{1}{\det A}(A(adj\, A)) = I_n$$

Según el teorema 3.2 inciso (4) (capítulo 1), esta última fórmula se puede escribir como

$$A\left(\frac{1}{\det A}\, adj\, A\right) = I_n$$

de donde

$$A^{-1} = \frac{1}{det\,A}\,adj\,A$$

Q.E.D.

En el capítulo 1(4.3) se desarrolló un *método* para encontrar la inversa de una matriz (basado en el proceso de eliminación Gaussiana). Véase, sin embargo, que el resultado del corolario anterior no es solamente "otro método" para encontrar la inversa. En él se está dando una *fórmula explícita* para A^{-1} en términos de A, la cual permite ver la "estructura interna" de la inversa de una matriz, información a la que no se tiene acceso con el método del capítulo 1. Se debe aclarar que la importancia de este corolario no es de carácter práctico: en general se tienen que hacer un mayor número de operaciones para hallar A^{-1} con la fórmula establecida que con el método del capítulo 1. Su importancia es solamente de carácter teórico (y estético, claro está).

EJEMPLO 3 A manera de ejemplo de aplicación de la fórmula (5.3) considere la matriz

$$A = \begin{bmatrix} 2 & 1 & 5 \\ 3 & -2 & 4 \\ 7 & 1 & 2 \end{bmatrix}$$

Encuentre los cofactores de los elementos de esta matriz

$$C_{11} = det\,A(1,1) = det\begin{bmatrix} -2 & 4 \\ 1 & 2 \end{bmatrix} = -8$$

$$C_{12} = det\,A(1,2) = -det\begin{bmatrix} 3 & 4 \\ 7 & 2 \end{bmatrix} = 22$$

$$C_{13} = det\,A(1,3) = det\begin{bmatrix} 3 & -2 \\ 7 & 1 \end{bmatrix} = 17$$

Similarmente encuentre

$C_{21} = 3$ $\quad$ $C_{31} = 14$

$C_{22} = -31$ $\quad$ $C_{32} = 7$

$C_{23} = 5$ $\quad$ $C_{33} = -7$

Entonces la matriz adjunta de A es:

$$adj\,A = \begin{bmatrix} -8 & 3 & 14 \\ 22 & -31 & 7 \\ 17 & 5 & -7 \end{bmatrix}$$

Se puede calcular *det A* desarrollando por cofactores con respecto a la primera línea para obtener

$$det\,A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = (2)(-8) + (1)(22) + (5)(17) = 91$$

Entonces

$$A^{-1} = \frac{1}{det\,A}\,\text{adj}\,A = \frac{1}{91}\begin{bmatrix} -8 & 3 & 14 \\ 22 & -31 & 7 \\ 17 & 5 & -7 \end{bmatrix}$$

5.2. LA REGLA DE CRAMER

Considérese el sistema de n ecuaciones con n incógnitas

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \end{aligned} \tag{5.4}$$

o bien, escrito matricialmente, $AX = B$, con $A = (a_{ij})_{i,j=1,\ldots,n}$, X la matriz $n \times 1$ de elementos $x_i(i=1,\ldots,n)$ y B la matriz $n \times 1$ de elementos $b_i(i=1,\ldots,n)$.

Supóngase que la matriz A es inversible. El teorema 4.10 (capítulo 1) dice entonces que el sistema (5.4) tiene solución única dada por

$$X = A^{-1}B$$

Al usar la fórmula (5.3) se puede escribir esta solución como

$$X = \frac{1}{det\,A}(adj\,A)(B)$$

Realícese el producto $(adj\,A)(B)$. El resultado será una matriz $n \times 1$ de elementos α_i, donde α_i se obtiene multiplicando la i-ésima línea de *adj A* (= i-ésima columna de la matriz de cofactores de A) por la matriz B; es decir,

$$\alpha_i = C_{1i}b_1 + C_{2i}b_2 + \ldots + C_{ni}b_n \tag{5.5}$$

Considérese la matriz A_i de orden n cuyos elementos coinciden con los correspondientes de A, excepto en la i-ésima columna, en donde se han colocado los elementos de la matriz B. Es decir,

$$A_i = \begin{bmatrix} a_{11} \ldots a_{1,i-1} & b_1 & a_{1,i+1} \ldots a_{1n} \\ a_{21} \ldots a_{2,i-1} & b_2 & a_{2,i+1} \ldots a_{2n} \\ & \cdots & \\ a_{n1} \ldots a_{n,i-1} & b_n & a_{n,i+1} \ldots a_{nm} \end{bmatrix}$$

$$\uparrow$$

i-ésima columna = matriz B

Calcúlese $det\ A_i$ desarrollando por cofactores respecto de la i-ésima columna de A_i.

Obsérvese que los cofactores involucrados serán los mismos cofactores correspondientes a la matriz A (pues para su cálculo no se emplea la i-ésima columna de A, que es en la única que difiere de A'). Entonces

$$det\ A_i = C_{1i}b_1 + C_{2i}b_2 + \ldots + C_{ni}b_n$$

Al comparar con la fórmula (5.5) se ve pues que $\alpha_i = det\ A_i$, por lo que la solución del sistema es:

$$X = \frac{1}{det\ A}(det\ A_i)_{i=1,\ldots,n}$$

o más explícitamente

$$x_i = \frac{det\ A_i}{det\ A} \qquad i = 1, \ldots, n \tag{5.6}$$

Las fórmulas (5.6) dan la solución explícita del sistema (5.4). Véase que ellas son aplicables sólo en el caso en el que la matriz del sistema sea inversible (esto es $det\ A \neq 0$). Se tiene en ese caso el siguiente método para resolver el sistema $AX = B$: se calcula el determinante de A (que tiene que ser distinto de cero para poder aplicar el método), así como también los determinantes de las matrices A_i, las cuales se obtienen de A sustituyendo en ésta los elementos de su i-ésima columna por los elementos de la matriz de términos independientes B. La solución del sistema es $x_i = \frac{det\ A_i}{det\ A}$ $i = 1, \ldots, n$. A este método se le conoce como *regla de Cramer*.

EJEMPLO 4 Suponga, por ejemplo, que se quiere resolver el sistema

$$\begin{aligned} x_1 + 2x_2 + 4x_3 &= 31 \\ 5x_1 + x_2 + 2x_3 &= 29 \\ 3x_1 - x_2 + x_3 &= 10 \end{aligned}$$

Se tiene

$$det\ A = det \begin{bmatrix} 1 & 2 & 4 \\ 5 & 1 & 2 \\ 3 & -1 & 1 \end{bmatrix} = -27$$

$$det\ A_1 = det \begin{bmatrix} 31 & 2 & 4 \\ 29 & 1 & 2 \\ 10 & -1 & 1 \end{bmatrix} = -81$$

$$det\ A_2 = det \begin{bmatrix} 1 & 31 & 4 \\ 5 & 29 & 2 \\ 3 & 10 & 1 \end{bmatrix} = -108$$

$$det\ A_3 = det \begin{bmatrix} 1 & 2 & 31 \\ 5 & 1 & 29 \\ 3 & -1 & 10 \end{bmatrix} = -135$$

Entonces

$$x_1 = \frac{det\ A_1}{det\ A} = \frac{-81}{-27} = 3,\ x_2 = \frac{det\ A_2}{det\ A} = \frac{-108}{-27} = 4,\ x_3 = \frac{det\ A_3}{det\ A} = \frac{-135}{-27} = 5$$

Se quiere hacer hincapié nuevamente en el hecho de que al resolver sistemas de ecuaciones, la regla de Cramer (cuando ésta puede aplicarse) resulta en general, desde el punto de vista computacional, un método menos eficiente que el método de eliminación Gaussiana descrito en el capítulo 1. No se puede negar, sin embargo, el gran valor estético de las fórmulas (5.6).

EJERCICIOS (SECCIÓN 5, CAPÍTULO 2)

1. Sean A y B las matrices

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 4 \\ 2 & 1 & -2 \end{bmatrix} \qquad B = \begin{bmatrix} -2 & 3 & 1 \\ 4 & 2 & 8 \\ 7 & -1 & 3 \end{bmatrix}$$

a) Determine la matriz adjunta de A y la matriz adjunta de B.
b) Calcule el producto $(adj\ B)(adj\ A)$.
c) Encuentre la matriz adjunta del producto AB (compare el resultado con el del inciso anterior).

② 2. ¿Cierto o falso? Toda matriz cuadrada de A conmuta con su matriz adjunta.

3. Determine la adjunta de la matriz

$$\begin{bmatrix} 1 & 3 & 5 \\ 0 & 7 & 8 \\ 0 & 0 & 9 \end{bmatrix}$$

4. Demuestre que la matriz adjunta de una matriz triangular superior (inferior) es una matriz triangular superior (inferior, respectivamente).
5. Compruebe que la matriz adjunta de una matriz diagonal es una matriz diagonal.
6. Sea A una matriz involutoria (véase ejercicio 26 de la sección 3, capítulo 1). Determine la matriz $adj\ A$.
7. Sea A una matriz idempotente (véase ejercicio 13 de la sección 3, capítulo 1). Suponga que A no es la matriz identidad. Demuestre que $A(adj\ A) = 0$. (*Sugerencia*: véase capítulo 27 de la sección 4, capítulo 1.)
8. Sea A una matriz nilpotente (véase ejercicio 14 de la sección 3, capítulo 1). Compruebe que $A(adj\ A) = 0$.

② 9. Demuestre que si el determinante de la matriz A de orden n ($n \geq 2$) es distinto de cero, entonces

$$det\ (adj\ A) = (det\ A)^{n-1}$$

10. Compruebe que si A es una matriz inversible, su matriz adjunta es también inversible. ¿Es válida la afirmación recíproca?
11. Sea A la matriz

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Demuestre que $adj(adj\ A) = A$.

③ 12. Sea A una matriz cuadrada de orden n ($n \geq 2$). Pruebe que si $det\ A \neq 0$, en ese caso

$$adj\ (adj\ A) = (det\ A)^{n-2}A$$

Concluya entonces que

$$det\ (adj\ (adj\ A)) = (det\ A)^{(n-1)^2}$$

13. Determine la inversa de las siguientes matrices, usando la fórmula (5.3)

a) $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 2 \\ 3 & 1 & 4 \end{bmatrix}$ b) $\begin{bmatrix} 2 & 1 & 3 \\ 5 & 7 & 1 \\ -1 & 1 & 4 \end{bmatrix}$ c) $\begin{bmatrix} 2 & 2 & 1 \\ 3 & -1 & -1 \\ 8 & 0 & 3 \end{bmatrix}$

14. Sea A una matriz de orden n con elementos enteros tal que $det\ A = 1$. Demuestre que su inversa tiene todos sus elementos enteros.

① 15. Sea A una matriz cuadrada de orden n y sea C una matriz de orden $n \times 1$. Denote por $A(j,C)$ la matriz (de orden n) cuyos elementos coinciden con los de A excepto en la columna j, en donde la matriz $A(j,C)$ tiene los elementos de la matriz C. Por ejemplo si

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 20 & 5 & 4 \\ -1 & 2 & 1 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 10 \\ 30 \\ -2 \end{bmatrix}$$

se tiene

$$A(1, C) = \begin{bmatrix} 10 & 3 & 8 \\ 30 & 5 & 4 \\ -2 & 2 & 1 \end{bmatrix} \quad A(2, C) = \begin{bmatrix} 1 & 10 & 8 \\ 20 & 30 & 4 \\ -1 & -2 & 1 \end{bmatrix} \quad A(3, C) = \begin{bmatrix} 1 & 3 & 10 \\ 20 & 5 & 20 \\ -1 & 2 & -2 \end{bmatrix}$$

Considere ahora el sistema de n ecuaciones con n incógnitas $AX = B$.

a) Demuestre que $(A)(I_n(j,X)) = A(j,B)$.
b) Compruebe que $det\ I_n(j,X) = x_j$ (la j-ésima incógnita del sistema).
c) Concluya de los dos incisos anteriores, que $(det\ A)\ x_j = det\ (A(j,B))$ y entonces, si $det\ A \neq 0$ se tiene

$$x_j = \frac{det\ A(j, B)}{det\ A} \qquad j = 1, 2, \ldots, n$$

Ésta es una nueva demostración de la regla de Cramer.

16. Al usar la regla de Cramer, dé una nueva demostración del teorema 4.3 del capítulo 1: si A es una matriz inversible, en ese caso el sistema homogéneo de ecuaciones lineales $AX = 0$ tiene sólo la solución trivial.
17. Al usar la regla de Cramer, resuelva los siguientes sistemas de ecuaciones:

a) $2x_1 + 3x_2 = 1$
$4x_1 - 2x_2 = 2$

b) $x_1 - 7x_2 = 4$
$8x_1 + 4x_2 = 3$

c) $3x_1 - 5x_2 + 2x_3 = -4$
$2x_1 - x_2 + 3x_3 = 9$
$4x_1 - 7x_2 + x_3 = 5$

d) $4x_1 - 3x_2 + 3x_3 = 8$
$x_1 + 3x_2 + x_3 = 7$
$3x_1 - x_2 + 2x_3 = 1$

e) $x_1 + 2x_2 + 2x_3 = 4$
$2x_1 + x_2 - 2x_3 = -1$
$2x_1 - 2x_2 + x_3 = -6$

④ 18. En este ejercicio se obtendrán algunos otros resultados interesantes sobre la función determinante, consecuencia de los cuales será la fórmula para la adjunta de un producto de matrices $adj(AB) = (adj\ B)(adj\ A)$, en donde A y B son dos matrices cuadradas del mismo orden.

Sean A y B dos matrices de orden n, $A = (a_{ij})_{i,j=1,\ldots,n}$, $B = (b_{ij})_{i,j=1,\ldots,n}$. Introduzca la notación $A(\ ,j)$ para la matriz de orden $n \times (n-1)$ que proviene de la matriz A suprimiendo en ésta su j-ésima columna. Similarmente, $A(i,\ )$ denotará la matriz de orden $(n-1) \times n$ que proviene de A suprimiendo en ésta su i-ésima línea.

En particular, la matriz $A(\ ,j)(i,\ ) = A(i,\ )(\ ,j)$ es la matriz de orden $(n-1) \times (n-1)$ que proviene de A eliminando en ésta su i-ésima línea y su j-ésima columna. Denote a esta matriz como $A(i,j)$. Recuerde también la definición de determinante de la matriz A.

$$det A = \sum_{\sigma} (sgn\ \sigma) a_{1\,\sigma(1)}\, a_{2\,\sigma(2)} \ldots a_{n\,\sigma(n)}$$

que escribirá como

$$det\ A = \sum_{\sigma \in S_n} (sgn\ \sigma) \prod_{r=1}^{n} a_{r\,\sigma(r)}$$

en donde la suma se hace sobre las $n!$ permutaciones del conjunto $S = \{1, 2, \ldots, n\}$ (al conjunto de todas estas permutaciones se está denotando por S_n).

a) Considere las matrices

$$A = \begin{bmatrix} 1 & 4 & 3 \\ 2 & 1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 7 & -1 \end{bmatrix}$$

a.1) Determine el producto AB.

a.2) Calcule $det\, AB$.

a.3) Encuentre las matrices $A(\ ,1)$, $A(\ ,2)$, $A(\ ,3)$, $B(1,\ )$, $B(2,\ )$ y $B(3,\ )$.

a.4) Determine los productos $A(\ ,1)B(1,\ )$, $A(\ ,2)B(2,\ )$ y $A(\ ,3)B(3,\ )$.

a.5.) Verifique que

$$det\, AB = det\, A(\ ,1)B(1,\ ) + det\, A(\ ,2)B(2,\ ) + det\, A(\ ,3)B(3,\ )$$

Éste es un resultado general que se probará en el siguiente inciso:

b) Considere las matrices $A = (a_{ij})_{\substack{i=1,\ldots,n-1\\ j=1,\ldots,n}}$ y $B = (b_{ij})_{\substack{i=1,\ldots,n\\ j=1,\ldots,n-1}}$. El producto AB es entonces una matriz cuadrada de orden $(n-1)$. Se quiere establecer el resultado

$$det\, AB = \sum_{s=1}^{n} det\, A(\ ,s)B(s,\ )$$

b.1) Demuestre que el resultado para el caso $n = 2$.

b.2) Compruebe que para $n \geq 3$

$$det\, AB = \sum_{\sigma \in S_{n-1}} (sgn\ \sigma) \left[\prod_{r=1}^{n-1} \left(\sum_{k=1}^{n} a_{rk} b_{k\,\sigma(r)} \right) \right]$$

b.3) Demuestre que para $n \geq 3$

$$\sum_{\sigma \in S_{n-1}} (sgn\ \sigma) \left[\sum_{k=1}^{n} \left(\prod_{r=1}^{n-1} a_{rk} b_{k\,\sigma(r)} \right) \right] = 0$$

b.4) Pruebe que para cualquier j y para $n \geq 3$

$$\prod_{r=1}^{n-1} \left(\sum_{k=1}^{n} a_{rk} b_{kj} \right) + \sum_{k=1}^{n} \left(\prod_{r=1}^{n-1} a_{rk} b_{kj} \right) = \sum_{s=1}^{n} \left[\prod_{r=1}^{n-1} \left(\sum_{\substack{k=1\\ k\neq s}}^{n} a_{rk} b_{kj} \right) \right]$$

b.5) Use los resultados de los tres subincisos anteriores para demostrar que

$$det\, AB = \sum_{s=1}^{n} \left\{ \sum_{\sigma \in S_{n-1}} (sgn\ \sigma) \left[\prod_{r=1}^{n-1} \left(\sum_{\substack{k=1\\ k\neq s}}^{n} a_{rk} b_{k\,\sigma(r)} \right) \right] \right\}$$

y concluya entonces que

$$det\, AB = \sum_{s=1}^{n} det\, A(\ ,s)B(s,\ )$$

c) Sean ahora A y B dos matrices de orden n. Demuestre que

$$A(i,\ )B(\ ,j) = AB(i,j)$$

d) En base a los resultados de los dos incisos anteriores, demuestre que siendo A y B matrices de orden n se tiene

$$det\, AB(i,j) = \sum_{k=1}^{n} det\, A(i,k)B(k,j)$$

Ejemplifique este resultado con las matrices

$$A = \begin{bmatrix} 10 & 8 & 5 \\ 1 & 4 & 3 \\ 2 & 1 & 1 \end{bmatrix} \qquad A = \begin{bmatrix} 1 & -2 & 8 \\ 3 & 4 & 7 \\ 7 & -1 & 2 \end{bmatrix}$$

con

d.1) $i = 1$ y $j = 3$ (véase el inciso a))

d.2) $i = 2$ y $j = 1$

d.3) $i = 3$ y $j = 2$

e) Considere ahora las matrices $adj\, A$ y $adj\, B$, adjuntas de las matrices A y B de orden n. Se quiere ver que $adj\,(AB) = (adj\, B)(adj\, A)$.

e.1) Sean A_{ij}, B_{ij} y AB_{ij} los cofactores asociados a los elementos a_{ij}, b_{ij} y $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$ de las matrices A, B y AB, respectivamente. Compruebe que el elemento de la i-ésima línea y j-ésima columna de la matriz $(adj\, B)(adj\, A)$ es:

$$\sum_{k=1}^{n} B_{ki}A_{jk}$$

el cual puede escribirse como

$$(-1)^{j+i} \sum_{k=1}^{n} det\, A(j,k)B(k,i)$$

e.2) Use el resultado de e.1) y del inciso d) para concluir que el elemento de la i-ésima línea y j-ésima columna de la matriz $(adj\, B)(adj\, A)$ es:

$$(-1)^{i+j}\, det\, AB(j,i) = AB_{ji}$$

e.3) Concluya, entonces, que $(adj\, B)(adj\, A) = adj\, AB$.

f) Use la fórmula (5.3) para demostrar nuevamente la validez de la fórmula $adj\, AB = (adj\, B)(adj\, A)$ en donde A y B son dos matrices del mismo orden *inversibles*.

APÉNDICE. EL DETERMINANTE DE VANDERMONDE

Un determinante que suele aparecer en relación a varios problemas en la matemática (por ejemplo, problemas de interpolación) es el llamado determinante de Vandermonde, que está dado por

$$V_n = det\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & \dots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \dots & x_n^2 \\ & & \dots & & \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \dots & x_n^{n-1} \end{bmatrix}$$

en donde $x_1, x_2, \dots, x_n$ son números distintos dados.

Por ejemplo si $n = 2$ se tiene

$$V_2 = det\begin{bmatrix} 1 & 1 \\ x_1 & x_2 \end{bmatrix}$$

$$= x_2 - x_1$$

Véase también el caso $n = 3$. Se tiene

$$V_3 = det\begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{bmatrix}$$

Se puede reducir el cálculo de V_3 al de un determinante de una matriz de orden 2 haciendo ceros las posiciones debajo del 1 en la primera columna y desarrollando por cofactores con respecto de ella

$$V_3 = det\begin{bmatrix} 1 & 1 & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 \\ 0 & x_2^2 - x_1^2 & x_3^2 - x_1^2 \end{bmatrix} = det\begin{bmatrix} x_2 - x_1 & x_3 - x_1 \\ x_2^2 - x_1^2 & x_3^2 - x_1^2 \end{bmatrix}$$

$$= (x_2 - x_1)(x_3^2 - x_1^2) - (x_2^2 - x_1^2)(x_3 - x_1) = (x_3 - x_2)(x_3 - x_1)(x_2 - x_1)$$

En el caso general, se afirma que

$$V_n = \prod_{n \geq i > j \geq 1} (x_i - x_j) \qquad \text{(A1)}$$

Es decir, V_n es el producto de todos los factores del tipo $(x_i - x_j)$ con $i > j$. Por ejemplo, en el caso $n = 2$ sólo se tiene el factor correspondiente a $2 > 1$, esto es $x_2 - x_1$. En el caso $n = 3$ se tienen 3 factores correspondientes a $3 > 2, 3 > 1, 2 > 1$, esto es $(x_3 - x_2)(x_3 - x_1)(x_2 - x_1)$.

En el caso $n = 4$ se tendrían 6 factores correspondientes a $4 > 3, 4 > 2, 4 > 1, 3 > 2, 3 > 1, 2 > 1$, esto es

$$V_4 = (x_4 - x_3)(x_4 - x_2)(x_4 - x_1)(x_3 - x_2)(x_3 - x_1)(x_2 - x_1)$$

Obsérvese que

$$V_3 = (x_3 - x_2)(x_3 - x_1)V_2 \qquad \text{(A2)}$$
$$V_4 = (x_4 - x_3)(x_4 - x_2)(x_4 - x_1)V_3$$

Para demostrar en general la fórmula (A1) se va a proceder por inducción.

Se ha visto ya su validez para $n = 2$ y $n = 3$.

Supóngase entonces que tal fórmula es válida para $n = k$; o sea que

$$V_k = \prod_{k \geq i > j \geq 1} (x_i - x_j)$$

y pruébese que también es válida para $n = k + 1$.

Obsérvese primeramente, que la diferencia entre V_{k+1} y V_k son los factores

$$(x_{k+1} - x_k)(x_{k+1} - x_{k-1}) \dots (x_{k+1} - x_1) = \prod_{k+1 > j \geq 1} (x_{k+1} - x_j)$$

como puede verse en los casos concretos $n = 2, 3$ y 4 en la fórmula (A2).

Es decir, V_{k+1} y V_k están relacionadas por

$$V_{k+1} = (x_{k+1} - x_k)(x_{k+1} - x_{k-1}) \dots (x_{k+1} - x_1)V_k$$
$$= \prod_{k+1 > j \geq 1} (x_{k+1} - x_j) \prod_{k \geq i > j \geq 1} (x_i - x_j) \qquad \text{(A3)}$$
$$= \prod_{k+1 \geq i > j \geq 1} (x_i - x_j)$$

Se procederá entonces a la demostración de la tesis de inducción (es decir, la fórmula (A3)).

Considérese entonces V_{k+1}

$$V_{k+1} = det\begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ x_1 & x_2 & \dots & x_k & x_{k+1} \\ x_1^2 & x_2^2 & \dots & x_k^2 & x_{k+1}^2 \\ & & \dots & & \\ x_1^k & x_2^k & \dots & x_k^k & x_{k+1}^k \end{bmatrix}$$

Realícense las siguientes operaciones elementales en las líneas de la matriz: cada línea va a sustituirse por ella misma menos x_{k+1} veces su línea anterior (comenzando por la última línea y yendo hacia arriba). El valor de V_{k+1} no se altera y así se obtiene

$$V_{k+1} = det \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ x_1 - x_{k+1} & x_2 - x_{k+1} & \dots & x_k - x_{k+1} & 0 \\ x_1(x_1 - x_{k+1}) & x_2(x_2 - x_{k+1}) & \dots & x_k(x_k - x_{k+1}) & 0 \\ \vdots & & & & \\ x_1^{k-1}(x_1 - x_{k+1}) & x_2^{k-1}(x_2 - x_{k+1}) & \dots & x_1^{k-1}(x_k - x_{k+1}) & 0 \end{bmatrix}$$

Al desarrollar por cofactores respecto de la última columna se obtiene (el signo del cofactor es $(-1)^{k+2} = (-1)^k$)

$$V_{k+1} = (-1)^k \; det \begin{bmatrix} x_1 - x_{k+1} & x_2 - x_{k+1} & \dots & x_k - x_{k+1} \\ x_1(x_1 - x_{k+1}) & x_2(x_2 - x_{k+1}) & \dots & x_k(x_k - x_{k+1}) \\ & & \dots & \\ x_1^{k-1}(x_1 - x_{k+1}) & x_2^{k-1}(x_2 - x_{k+1}) & \dots & x_1^{k-1}(x_k - x_{k+1}) \end{bmatrix}$$

Multiplíquese la j-ésima columna de esta matriz por $\dfrac{1}{x_j - x_{k+1}}$ $(j = 1, \dots, k)$. Obsérvese que el valor del determinante quedará alterado por los factores $(x_j - x_{k+1})(j = 1, 2, \dots, k)$. Queda entonces

$$V_{k+1} = (-1)^k(x_1 - x_{k+1})(x_2 - x_{k+1}) \dots (x_k - x_{k+1}) \; det \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_k \\ & \dots & & \\ x_1^{k-1} & x_2^{k-1} & \dots & x_k^{k-1} \end{bmatrix}$$

o sea,

$$V_{k+1} = (x_{k+1} - x_k) \dots (x_{k+1} - x_2)(x_{k+1} - x_1) V_k$$

relación que se quería establecer.

EJERCICIOS (APÉNDICE, CAPÍTULO 2)

1. Demuestre que $V_n = 0$ si, y sólo si $x_i = x_j$ para algún $i \neq j$, $i, j \in \{1, 2, \dots, n\}$.
2. Calcule los siguientes determinantes:

a) $det \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$

b) $det \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & -2 \\ 1 & 1 & 4 & 4 \\ 1 & -1 & 8 & -8 \end{bmatrix}$

c) $det \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 3 & -3 \\ 1 & 0 & 1 & 9 & 9 \\ -1 & 0 & 1 & 27 & -27 \\ 1 & 0 & 1 & 81 & 81 \end{bmatrix}$

3. Sean x_1, x_2, x_3 tres números reales distintos. Compruebe que existe un único polinomio $p(x)$ de grado 2 $p(x) = ax^2 + bx + c$, tal que $p(x_1) = y_1$, $p(x_2) = y_2$, $p(x_3) = y_3$ en donde y_1, y_2, y_3 son números dados.

   (*Sugerencia*: demuestre que el sistema de ecuaciones $p(x_i) = y_i$, $i = 1, 2, 3$ tiene una única solución para los coeficientes a, b, c.)

② 4. Sean $x_1, x_2, \dots, x_{n+1}$ $n + 1$ números reales distintos. Demuestre que existe un único polinomio $p(x)$ de grado n

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0$$

tal que $p(x_i) = y_i$, $i = 1, 2, \dots, n + 1$, en donde $y_1, y_2, \dots, y_{n+1}$ son números dados. Pruebe también que si el polinomio $p(x)$ se anula en $x_1, x_2, \dots, x_{n+1}$ entonces $p(x)$ es el polinomio idénticamente nulo.

CAPÍTULO TRES

Espacios vectoriales

Todos conocen las siguientes propiedades de los números reales (pues en ellos descansa la manipulación algebraica que se ha aprendido desde la secundaria): si a, b, c son números reales se tiene

1) $a + b = b + a$
2) $a + (b + c) = (a + b) + c$
3) $a + 0 = a$
4) $a + (-a) = 0$
5) $ab = ba$
6) $a(bc) = (ab)c$
7) $a \cdot 1 = a$
8) si $a \neq 0$, $(a)(a^{-1}) = 1$
9) $a(b + c) = ab + ac$

Considérese el conjunto $A = \{0, 1, 2\}$ y en él defínanse las operaciones de suma y producto entre sus elementos según las siguientes "tablas de suma y multiplicación"

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Se puede ver que las nueve propiedades de los números reales mencionadas anteriormente son también válidas para este conjunto A con estas operaciones de suma y de producto. En efecto

1) Si $a, b \in A$, se tiene siempre $a + b = b + a$. Por ejemplo, $1 + 2 = 0 = 2 + 1$.
2) Si $a, b, c \in A$, se tiene siempre $a + (b + c) = (a + b) + c$. Por ejemplo, $1 + (2 + 2) = 1 + 1 = 2$ y $(1 + 2) + 2 = 0 + 2 = 2$, de modo que $1 + (2 + 2) = (1 + 2) + 2$.
3) Para cualquier $a \in A$, se tiene siempre $a + 0 = a$.

4) Dado un elemento $a \in A$, existe siempre otro elemento $(-a) \in A$ (el inverso aditivo de A) tal que $a + (-a) = 0$. De hecho, $(-0) = 0$, $(-1) = 2$, $(-2) = 1$.
5) Si $a, b \in A$, se tiene siempre $ab = ba$.
6) Si $a, b, c \in A$ se tiene siempre $a(bc) = (ab)c$. Por ejemplo, $2(2 \cdot 1) = 2(2) = 1$ y $(2 \cdot 2) \cdot 1 = (1) \cdot 1 = 1$, de modo que $2 \cdot (2 \cdot 1) = (2 \cdot 2) \cdot 1$.
7) Para cualquier $a \in A$ se tiene siempre $a \cdot 1 = a$.
8) Dado un elemento $a \in A$, $a \neq 0$, existe siempre un elemento $(a^{-1}) \in A$ (el inverso multiplicativo de a) tal que $a(a^{-1}) = 1$. De hecho, $(1^{-1}) = 1$, $(2^{-1}) = 2$.
9) Si $a, b, c \in A$ se tiene siempre $a(b + c) = ab + ac$. Por ejemplo, $2(2 + 1) = 2(0) = 0$ y $2 \cdot 2 + 2 \cdot 1 = 1 + 2 = 0$, de modo que $2(2 + 1) = 2 \cdot 2 + 2 \cdot 1$.

De modo, pues, que el conjunto $A = \{0, 1, 2\}$ con las operaciones de suma y producto dadas, tiene las mismas propiedades (1) - (9) de los números reales.

Existen muchos otros ejemplos interesantes de conjuntos en los que se tienen definidas las operaciones de suma y producto, que también satisfacen las propiedades (1) - (9) de los reales. Por ejemplo:

a) El conjunto de números racionales $Q = \left\{ \frac{m}{n} \,\middle|\, m, n \in \mathbf{Z}, n \neq 0 \right\}$ con las operaciones usuales.

$$\frac{m_1}{n_1} + \frac{m_2}{n_2} + \frac{m_1 n_2 + m_2 n_1}{n_1 n_2}$$

$$\left(\frac{m_1}{n_1}\right)\left(\frac{m_2}{n_2}\right) = \frac{m_1 m_2}{n_1 n_2}$$

b) El conjunto de números complejos $C = \{a + ib \mid a, b \in \mathbf{R}\}$ con las operaciones usuales

$$(a + ib) + (c + id) = (a + c) + i(b + d)$$

$$(a + ib)(c + id) = (ac - bd) + i(ad + bc)$$

entre otros.

Un conjunto no vacío en el que existen las operaciones de suma y producto satisfaciendo las propiedades (1) - (9) enunciadas en la página anterior se le llama *CAMPO*.

Un campo es, pues, una estructura algebraica, la cual es compartida por **R**, **Q**, **C** y nuestro conjunto $A = \{0, 1, 2\}$.

Ante esta situación, la dinámica natural de la matemática lleva de inmediato al siguiente planteamiento: ¿qué se puede decir en general sobre un campo? Dentro de algunas partes del álgebra (abstracta) se da respuesta a esta pregunta.

Una situación completamente análoga a la planteada anteriormente es la que se comenzará a abordar en este capítulo; se estudiará (en abstracto) una nueva estructura algebraica: la de espacio vectorial.

En los cursos de física se ha aprendido a sumar vectores en el plano (o en el espacio), así como a multiplicarlos por un número real. Estas operaciones entre vectores y reales poseen ciertas propiedades que son compartidas por otros conjuntos en los que existen también ciertas operaciones de suma y producto por números reales. Tales conjuntos (con sus operaciones) se les llama en general *espacios vectoriales.*

La estructura algebraica de espacio vectorial es *la más importante dentro del álgebra lineal.*

Éste es, pues, el capítulo en el que se abre propiamente el desfile de ideas y resultados importantes del álgebra lineal. Se advierte al lector que el sabor de la abstracción matemática, que se mantuvo apenas insinuado en los dos capítulos anteriores, se dejará sentir más fuertemente a partir de este capítulo.

1. DEFINICIÓN Y OBSERVACIONES PRELIMINARES

En esta primera sección del capítulo, se definirá formalmente el concepto de espacio vectorial y se obtendrán algunas consecuencias inmediatas de tal definición.

Aclaración: Como se verá a continuación, una de las características que definen a un espacio vectorial es la operación llamada "multiplicación por escalares". En un acercamiento más abstracto a la teoría de espacios vectoriales, se pide solamente que estos escalares sean elementos de un campo K cualquiera. En tal caso se dice que se trata de un "espacio vectorial sobre el campo K". Por ejemplo, si K es el campo de los números reales, se dice que el espacio vectorial es "sobre los reales" o simplemente que es un *espacio vectorial real.* Similarmente, cuando K es el campo de los números complejos, se dice que se trata de un *espacio vectorial complejo*. En este libro no se será pretencioso en cuanto a la generalidad del campo de escalares. Se limitará uno a considerar el caso en el que tal campo es el de los números reales. Es decir, todos los espacios vectoriales serán espacios vectoriales reales. Por tanto, cuando se hable de "escalares" se estará uno refiriendo siempre a números reales (a menos que se diga explícitamente lo contrario).

DEFINICIÓN 1.1 Un espacio vectorial es un *conjunto* no vacío V en el cual están definidas *dos operaciones*

1) $+: V \times V \to V, (v_1, v_2) \to v_1+v_2$
2) $\cdot: R \times V \to V, (k, v) \to kv$*

*El hecho de que las funciones + y · tomen valores *en V*, indica la propiedad de *cerradura* de las operaciones en el espacio vectorial.

llamadas *suma* y *producto por escalares* (respectivamente) y las cuales satisfacen los siguientes axiomas:

1) $u + v = v + u,\ \forall\, u, v \in V$ (la suma es conmutativa)
2) $u + (v + w) = (u + v) + w,\ \forall\, u, v, w \in V$ (la suma es asociativa)
3) Existe un elemento $0 \in V$, llamado *cero,* con la propiedad
$$u + 0 = u \qquad \forall\, u \in V$$
4) Para cada $v \in V$, existe un elemento $(-v) \in V$, llamado "el inverso aditivo de v" con la propiedad
$$v + (-v) = 0$$
5) $\lambda(u + v) = \lambda u + \lambda v,\ \forall\, \lambda \in \mathbf{R}, \forall\, u, v \in V.$
6) $(\lambda + \mu)v = \lambda v + \mu v,\ \forall\, \lambda, \mu \in \mathbf{R}, \forall\, v \in V.$
7) $(\lambda\, \mu)v = \lambda(\mu v),\ \forall\, \lambda, \mu \in \mathbf{R}, \forall\, v \in V.$
8) $1 \cdot v = v,\ \forall\, v \in V.$

A los elementos de un espacio vectorial se les llama "vectores". A partir de este momento, entonces, la palabra "vector" no tendrá relación alguna con el concepto geométrico de "flecha" (relación que se encuentra naturalmente en el estudiante por su experiencia en cursos de física), aunque se verá que sí existen de hecho algunos "espacios vectoriales de flechas".

En vista de la generalidad de la definición de espacio vectorial, se podrán tener, entonces, espacios vectoriales de "objetos cualesquiera" (estos objetos serán los vectores) con tal de que en ese conjunto de objetos exista una suma y un producto por escalares, y que éstos verifiquen los axiomas requeridos.

En la siguiente sección se estudiarán varios ejemplos de espacios vectoriales importantes en matemáticas. El resto de la presente sección se dedicará a obtener algunas consecuencias de la definición de espacio vectorial establecida anteriormente.

Primeramente obsérvese que en un espacio vectorial V se tiene definida la operación de suma entre *dos vectores*. A cada par de vectores $u, v \in V$ le corresponde un nuevo vector $u + v \in V$ llamado "la suma de u y v". Se puede, sin embargo, extender esta operación a un número cualquiera (finito) de vectores $v_1, v_2, \ldots, v_n$ en el espacio vectorial V. Es decir, para un conjunto de n vectores $v_1, v_2, \ldots, v_n$ en el espacio vectorial V, se puede asociar un nuevo vector $v_1 + v_2 + \ldots + v_n \in V$, llamado "la suma de $v_1, v_2, \ldots, v_n$" de la siguiente manera: de la definición de espacio vectorial se sabe cuál es el vector $u_1 = v_1 + v_2$. Se define entonces la suma de v_1, v_2, v_3, como el vector $u_2 = u_1 + v_3$. Similarmente, la suma de v_1, v_2, v_3, v_4 será el vector $u_3 = u_2 + v_4$, etc. En general, una vez definido el vector $u_{n-2} \in V$, se define la suma de $v_1, v_2, \ldots, v_n$ como el vector $u_{n-1} = u_{n-2} + v_n$ ($n \geq 3$). (Se trata pues de una definición inductiva.) Además, en vista de la propiedad asociativa de la suma, este vector se puede denotar como $v_1 + v_2 + \ldots + v_n$, sin temor a confusión.

En el axioma (3) se establece la existencia de un elemento del espacio vectorial que actúa de manera neutra para la suma. Este elemento se llama "cero" y se denota

por 0. Es fácil ver que en un espacio vectorial solamente puede existir *un* cero. En efecto, supóngase que existen dos ceros en V, dígase 0_1 y 0_2. Entonces

$$\underset{Ax(3)}{0_1 = 0_1 + 0_2} = \underset{Ax(1)}{0_2 + 0_1} = \underset{Ax(3)}{0_2}$$

En el axioma (4) se establece la existencia de un inverso aditivo para cada elemento v del espacio vectorial. También es fácil ver que éste es *único*: supóngase que para $v \in V$ existen dos inversos aditivos $(-v)_1$ y $(-v)_2$. Entonces

$$(-v)_1 \underset{Ax(3)}{=} (-v)_1 + 0 \underset{Ax(4)}{=} (-v)_1 + (v+(-v)_2) \underset{Ax(2)}{=} ((-v)_1 + v) + (-v)_2 =$$

$$\underset{Ax(1)}{=} (v + (-v)_1) + (-v)_2 \underset{Ax(4)}{=} 0 + (-v)_2 \underset{Ax(1)}{=} (-v)_2 + 0 \underset{Ax(3)}{=} (-v)_2$$

En el siguiente teorema se establecen algunos hechos básicos (algunos de ellos son consecuencia inmediata de la definición de espacio vectorial) que son válidos para todos los espacios vectoriales.

TEOREMA 1.1

Sea V un espacio vectorial. Entonces

1) $0 \cdot v = 0 \quad \forall v \in V$
2) $(-1)v = -v \quad \forall v \in V$
3) $\lambda \cdot 0 = 0 \quad \forall \lambda \in R$
4) $-(-v) = v \quad \forall v \in V$
5) $(-\lambda)v = -(\lambda v) = \lambda(-v) \quad \forall \lambda \in R, \forall v \in V$
6) $(-\lambda)(-v) = \lambda v \quad \forall \lambda \in R, \quad \forall v \in V$
7) $u + w = v + w \Rightarrow u = v, \quad \forall u, v, w \in V$
8) Si $\lambda \in R$, y $\lambda \neq 0$, $\lambda u = \lambda v \Rightarrow u = v, \forall u, v \in V$
9) Si $\lambda v = 0 \Rightarrow \lambda = 0$ o $v = 0$.

DEMOSTRACIÓN La justificación de cada paso realizado en la demostración se basa en alguno (o algunos) de los axiomas que definen un espacio vectorial, o bien en alguno de los incisos previamente demostrado: déjase al lector que "complete" la demostración indicando cuál o cuáles hechos se están usando en cada signo de igualdad establecido.

1) $v + 0 \cdot v = 1 \cdot v + 0 \cdot v = (1 + 0) \cdot v = 1 \cdot v = v$
Entonces (unicidad del cero), $0 \cdot v = 0$.
2) $v + (-1)v = 1 \cdot v + (-1) \cdot v = (1+(-1)) \cdot v = 0 \cdot v = 0$
Entonces (unicidad del inverso aditivo), $(-1)v = -v$
3) $\lambda \cdot 0 = \lambda(0 \cdot v) = (\lambda \cdot 0) \cdot v = 0 \cdot v = 0$
4) $v + (-v) = (-v) + v = 0$. Entonces (unicidad del inverso aditivo), $-(-v) = v$.
5) $(-\lambda)v = ((-1)\lambda)\, v = (-1)(\lambda v) = -\lambda v$
$= (\lambda(-1))v = \lambda((-1)v) = \lambda(-v)$
6) $(-\lambda)(-v) = (-\lambda)((-1)v) = (-\lambda)(-1)v) = \lambda v$
7) $u + w = v + w \Rightarrow u + w + (-w) = v + w + (-w) \Rightarrow u + 0 = v + 0 \Rightarrow u = v$
8) $\lambda u = \lambda v \underset{(\lambda \neq 0)}{\Rightarrow} \lambda^{-1}(\lambda u) = \lambda^{-1}(\lambda v) \Rightarrow (\lambda^{-1}\lambda)u = (\lambda^{-1}\lambda)\, v \Rightarrow 1 \cdot u = 1 \cdot v \Rightarrow u = v$
9) Supóngase $\lambda \neq 0$. Entonces $\lambda v = 0 \Rightarrow \lambda^{-1}(\lambda v) = \lambda^{-1} 0 = 0 \Rightarrow (\lambda^{-1}\lambda)\, v = 0$
$\Rightarrow 1 \cdot v = 0 \Rightarrow v = 0$.

Q.E.D.

Obsérvese entonces que el inverso aditivo del vector $v \in V$ se obtiene multiplicando el vector v por el escalar -1 (inciso (2) del teorema anterior). También, nótese que en un espacio vectorial se tienen "leyes de cancelación" (incisos (7) y (8)).

Se desea hacer hincapié en el hecho de que un espacio vectorial *no es solamente un conjunto de objetos* (de vectores). Para que un conjunto sea un espacio vectorial, deben existir definidas en él operaciones de suma y producto por escalares, y aún más, estas operaciones deben de satisfacer los 8 axiomas de la definición de espacio vectorial.

Por ejemplo, considérese el conjunto de reales positivos

$$R^+ = \{x \in R \mid x > 0\}$$

Éste es, pues, sólo un conjunto.

Denótense a los elementos de este conjunto con las operaciones "usuales" de suma y producto por escalares (si $u, v, \in \mathbf{R}^+$, $u + v$ es la suma de los reales u y v y si $\lambda \in \mathbf{R}$ y $v \in \mathbf{R}^+$, λv es el producto del real λ por el real v). Este hecho no garantiza aún que se tenga ahora un espacio vectorial, pues si bien se tiene ya un conjunto y operaciones de suma y producto por escalares definidas en él, aún no se sabe si estas operaciones satisfacen los 8 axiomas de la definición de espacio vectorial. De hecho, es fácil ver que $\mathbf{R}^+$ con estas operaciones "usuales" *no* es un *espacio vectorial* (en principio, la operación de producto por escalares no es cerrada, en el sentido de que para cualquier $\lambda \in \mathbf{R}$ y $v \in \mathbf{R}^+$ se tenga $\lambda v \in \mathbf{R}^+$; tome por ejemplo $\lambda \in \mathbf{R}$, $\lambda < 0$).

Sin embargo, véase que este mismo conjunto $\mathbf{R}^+$, dotado con otras operaciones de suma y producto por escalares, puede convertirse en un espacio vectorial.

Defínase:

$u, v \in \mathbf{R}^+$, $u + v = uv$ (el producto usual de u y v)

$\lambda \in \mathbf{R}$, $v \in \mathbf{R}^+$, $\lambda v = v^\lambda$ (el real v elevado a la potencia λ)

(por ejemplo, si $u = 3$, $v = 9$, se tiene $u + v = 3 + 9 = (3)(9) = 27$, y si $\lambda = -2$ y $v = 4$, se tiene $\lambda v = (-2)(4) = (4)^{-2} = \frac{1}{16}$).

En principio, obsérvese que para $u, v \in \mathbf{R}^+$ se tiene $u + v = uv \in \mathbf{R}^+$ (el producto de reales positivos es un real positivo) y si $\lambda \in \mathbf{R}$ y $v \in \mathbf{R}^+$ se tiene $\lambda v = v^\lambda \in \mathbf{R}^+$ (al elevar un real positivo a cualquier potencia se obtiene un real positivo).

Al usar algunas propiedades elementales de los números reales así como algunas leyes de los exponentes, es fácil verificar que en este caso $\mathbf{R}^+$ *es* un espacio vectorial (el cero de este espacio es $1 \in \mathbf{R}^+$ y, para un vector $v \in \mathbf{R}^+$, su inverso aditivo $(-v)$ es $\frac{1}{v} \in \mathbf{R}^+$).

EJERCICIOS (SECCIÓN 1, CAPÍTULO 3)

1. Considere el conjunto

$$\mathbf{R}^2 = \{(x, y) \mid x, y \in \mathbf{R}\}$$

Defina las operaciones de suma y producto por escalares en este conjunto como

$$(x, y) + (x', y') = (x + x', y + y')$$
$$c(x, y) = (cx, cy) \qquad c \in \mathbf{R}$$

Demuestre que el conjunto $\mathbf{R}^2$ con estas operaciones es un espacio vectorial.

2. En el mismo conjunto $\mathbf{R}^2$ del ejercicio anterior, defina las operaciones de suma y producto por escalares como

$$(x, y) + (x', y') = (x - x', y - y')$$
$$c(x, y) = (cx, cy) \qquad c \in \mathbf{R}$$

¿Es $\mathbf{R}^2$ con estas operaciones un espacio vectorial?

3. Considere el conjunto

$$\mathbf{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbf{R}\}$$

En él se definen las operaciones de suma y producto por escalares como se indica en cada uno de los siguientes incisos. Determine en cada caso si el conjunto $\mathbf{R}^3$ con tales operaciones es un espacio vectorial.

a) $(x, y, z) + (x', y', z') = (x + x', y + y', z + z')$
$c(x, y, z) = (cx, cy, cz) \qquad c \in \mathbf{R}$

b) $(x, y, z) + (x', y', z') = (x + x', y + y', 0)$
$c(x, y, z) = (cx, cy, 0) \qquad c \in \mathbf{R}$

c) $(x, y, z) + (x', y', z') = (xx', yy', zz')$
$c(x, y, z) = (cx, cy, cz) \qquad c \in \mathbf{R}$

d) $(x, y, z) + (x', y', z') = (x + x', y + y', z + z')$
$c(x, y, z) = (5cx, 5cy, 5cz) \qquad c \in \mathbf{R}$

4. Considere el conjunto $M_{2\times 2}$ de todas las matrices cuadradas de orden 2

$$M_{2\times 2} = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a, b, c, d \in \mathbf{R} \right\}$$

Defina en $M_{2\times 2}$ la suma y el producto por escalares como

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} a + a' & b + b' \\ c + c' & d + d' \end{bmatrix}$$

$$k \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix}, \quad k \in \mathbf{R}$$

Demuestre que el conjunto $M_{2\times 2}$ con estas operaciones es un espacio vectorial.

① 5. En el mismo conjunto $M_{2\times 2}$ del ejercicio anterior defina la suma como

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} aa' + bc' & ab' + db' \\ ca' + dc' & cb' + dd' \end{bmatrix}$$

a) Demuestre que esta suma es asociativa.

b) Demuestre que existe un elemento neutro para esta suma, esto es, una matriz C de orden 2 tal que $A + C = A$ para toda A en $M_{2\times 2}$.

c) ¿Existe un inverso aditivo *para toda* matriz de este conjunto?

d) Demuestre que el conjunto $M_{2\times 2}$ con esta operación de suma y el mismo producto por escalares del ejercicio anterior no es un espacio vectorial.

6. Considere el conjunto Z cuyo único elemento es "zapato" (Z es un conjunto constituido por un zapato)

$$Z = \{\text{zapato}\}$$

En Z defina las operaciones de suma y producto por escalares como

$$\text{zapato} + \text{zapato} = \text{zapato}$$
$$k(\text{zapato}) = \text{zapato}, \qquad k \in \mathbf{R}$$

Demuestre que el conjunto Z con estas operaciones de suma y producto por escalares es un espacio vectorial.

① 7. En el ejercicio anterior se ha mostrado un ejemplo de un espacio vectorial con un solo elemento. ¿Existen espacios vectoriales con sólo dos elementos?, ¿con tres?, ¿con un número finito de elementos?

8. Sean $v_1, v_2, \ldots, v_n$ n vectores del espacio vectorial V y $c_1, c_2, \ldots, c_n$ n escalares cualesquiera. En base a la definición 1.1, establezca un sentido preciso para el vector $w \in V$ dado por

$$w = c_1v_1 + c_2v_2 + \ldots + c_nv_n$$

2. EJEMPLOS DE ESPACIOS VECTORIALES

Considérese el conjunto $\{v\}$ formado por un solo elemento v. En este conjunto defínanse las operaciones de suma y producto por escalares de la siguiente manera:

$$v + v = v$$

$$\lambda v = v, \qquad \lambda \in \mathbf{R}$$

Es fácil ver que el conjunto $\{v\}$ con estas operaciones forma un espacio vectorial.

En particular, por el axioma (3) de la definición de espacio vectorial, el elemento v debe ser el cero del espacio. Se escribirá entonces este espacio como $\{0\}$ con sus operaciones $0 + 0 = 0$, $\lambda 0 = 0$, $\lambda \in \mathbf{R}$. Éste es un espacio vectorial poco interesante. Se llamará *espacio vectorial trivial*. La importancia de este espacio es de carácter teórico.

El objetivo de esta sección es presentar algunos otros espacios vectoriales (más interesantes que el trivial) con los que se trabajará en el resto del libro.

2.1. EL ESPACIO $\mathbf{R}^n$

Éste es el ejemplo más importante de espacio vectorial, pues en cierto sentido que se precisará después (en la sección 6.2), muchos de los espacios vectoriales con los que se trabajará, son "equivalentes" a este espacio.

El conjunto de vectores en este espacio es el conjunto de n-adas ordenadas de números reales $x_1, x_2, \ldots, x_n$, es decir

$$\mathbf{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbf{R},\ i = 1, \ldots, n\}^*$$

Al vector $(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$ se le denotará simplemente como x. Es decir, se escribirá $x = (x_1, x_2, \ldots, x_n)$. A los números $x_1, x_2, \ldots, x_n$ se les llama *coordenadas* (o componentes) del vector x. Similarmente se escribirá $y = (y_1, y_2, \ldots, y_n)$, $z = (z_1, z_2, \ldots, z_n)$ etc., para denotar a los vectores de este espacio.

En $\mathbf{R}^n$ se define la suma de los vectores $x = (x_1, x_2, \ldots, x_n)$ y $y = (y_1, y_2, \ldots, y_n)$ coordenada a coordenada, es decir $x + y$ es el vector $\mathbf{R}^n$ cuyas coordenadas son las correspondientes sumas de coordenadas de x y y. O sea,

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

Similarmente, el producto del vector $x = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$ por el escalar $\lambda \in \mathbf{R}$ se define coordenada a coordenada, o sea,

$$\lambda x = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$$

*En este conjunto se define la igualdad entre sus elementos como:

$$(x_1, x_2, \ldots, x_n) = (x'_1, x'_2, \ldots, x'_n) \text{ si, y sólo si } x_i = x'_i,\ i = 1, \ldots, n.$$

Verifíquese que $\mathbf{R}^n$ con estas operaciones es efectivamente un espacio vectorial. Sean $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$, $z = (z_1, z_2, \ldots, z_n)$ tres vectores cualesquiera de $\mathbf{R}^n$ y λ, μ dos escalares. Se tiene

1) La suma es conmutativa

$$\begin{aligned} x + y &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \\ &= (y_1 + x_1, y_2 + x_2, \ldots, y_n + x_n \\ &= y + x \end{aligned}$$

(propiedad de conmutatividad de la suma de números reales)

2) La suma es asociativa

$$\begin{aligned} x + (y + z) &= (x_1, x_2, \ldots, x_n) + (y_1 + z_1, y_2 + z_2, \ldots, y_n + z_n) \\ &= (x_1 + (y_1 + z_1), x_2 + (y_2 + z_2), \ldots, x_n + (y_n + x_n)) \\ &= ((x_1 + y_1) + z_1, (x_2 + y_2) + z_2, \ldots, (x_n + y_n) + z_n) \\ &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) + (z_1, z_2, \ldots, z_n) \\ &= (x + y) + z. \end{aligned}$$

(propiedad asociativa de la suma de números reales)

3) Existe el cero en $\mathbf{R}^n$. Escríbase $0 \in \mathbf{R}^n$ como $0 = (0, 0, \ldots, 0)$. Entonces

$$\begin{aligned} x + 0 &= (x_1 + 0, x_2 + 0, \ldots, x_n + 0) \\ &= (x_1, x_2, \ldots, x_n) = x \end{aligned}$$

(propiedad neutra para la suma del cero en los números reales)

4) Para $x \in \mathbf{R}^n$, existe $(-x) \in \mathbf{R}^n$ tal que $x + (-x) = 0$. Escríbase

$$-x = (-x_1, -x_2, \ldots, -x_n) \in \mathbf{R}^n$$

Por lo tanto

$$x + (-x) = (x_1 - x_1, x_2 - x_2, \ldots, x_n - x_n) = (0, 0, \ldots, 0) = 0$$

5)
$$\begin{aligned} \lambda(x + y) &= \lambda(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \\ &= (\lambda(x_1 + y_1), \lambda(x_2 + y_2), \ldots, \lambda(x_n + y_n)) \\ &= (\lambda x_1 + \lambda y_1, \lambda x_2 + \lambda y_2, \ldots, \lambda x_n + \lambda y_n) \\ &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n) + (\lambda y_1, \lambda y_2, \ldots, \lambda y_n) \\ &= \lambda x + \lambda y. \end{aligned}$$

(propiedad distributiva de los números reales)

6) $(\lambda + \mu) x = ((\lambda + \mu)x_1, (\lambda + \mu)x_2, \ldots, (\lambda + \mu)x_n)$
$= (\lambda x_1 + \mu x_1, \lambda x_2 + \mu x_2, \ldots, \lambda x_n + \mu x_n)$ (propiedad distributiva de los números reales)
$= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n) + (\mu x_1, \mu x_2, \ldots, \mu x_n)$
$= \lambda x + \mu x.$

7) $(\lambda \mu)x = ((\lambda\mu)x_1, (\lambda\mu)x_2, \ldots, (\lambda\mu)x_n)$
$= (\lambda(\mu x_1), \lambda(\mu x_2), \ldots, \lambda(\mu x_n))$ (propiedad asociativa del producto de números reales)
$= \lambda(\mu x_1, \mu x_2, \ldots, \mu x_n)$
$= \lambda(\mu x)$

8) $1 \cdot x = (1 \cdot x_1, 1 \cdot x_2, \ldots, 1 \cdot x_n)$ (propiedad neutra para el producto del uno en los números reales)
$= (x_1, x_2, \ldots, x_n) = x$

Esto demuestra entonces que $\mathbf{R}^n$ es un espacio vectorial.

Debido a la importancia del carácter geométrico de este espacio en los casos $n = 2$ y $n = 3$, se hará una discusión particular para cada uno de estos dos casos.

I. El espacio $\mathbf{R}^2$

El conjunto $\mathbf{R}^2$ es entonces el conjunto de parejas ordenadas de números reales (x_1, x_2). Se escribirán en este caso las parejas como (x, y). Entonces

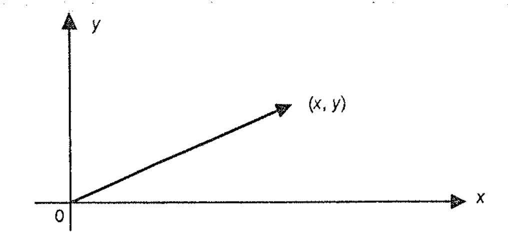
$$\mathbf{R}^2 = \{(x, y) \mid x, y \in \mathbf{R}\}$$

Las operaciones en $\mathbf{R}^2$ de suma y producto por escalares definidas por

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$
$$\lambda(x, y) = (\lambda x, \lambda y), \qquad \lambda \in \mathbf{R}$$

hacen de $\mathbf{R}^2$ un espacio vectorial.

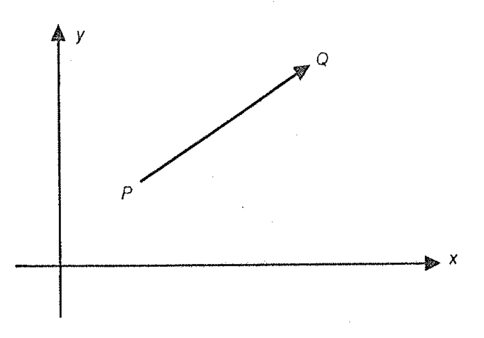
A cada vector $(x, y) \in \mathbf{R}^2$ se le puede identificar con un punto del plano xy. Aún más, para dar un sentido más geométrico a los elementos de este espacio vectorial, identifíquese al vector $(x, y) \in \mathbf{R}^2$ con el segmento dirigido que tiene su punto inicial en el origen y su punto final en el punto (x, y) como se muestra en la figura.



A partir de ahora no se hará distinción alguna entre "puntos del plano" y "segmentos dirigidos con inicio en el origen y final en el punto en cuestión".

Obsérvese que el vector cero de $\mathbf{R}^2$ es identificado con el origen del plano xy.

En este espacio vectorial se puede entonces ver a los vectores (a los elementos del espacio vectorial) como "vectores" (segmentos dirigidos) en el plano xy. Es decir, bajo esta perspectiva, todos los vectores del espacio vectorial $\mathbf{R}^2$ son segmentos dirigidos en el plano xy. ¿Será verdad la afirmación recíproca? Es decir, ¿un segmento dirigido en el plano xy es un vector de $\mathbf{R}^2$? Por ejemplo, se podría preguntar si el segmento dirigido $\overrightarrow{PQ}$ que se muestra en la figura es un vector de $\mathbf{R}^2$.



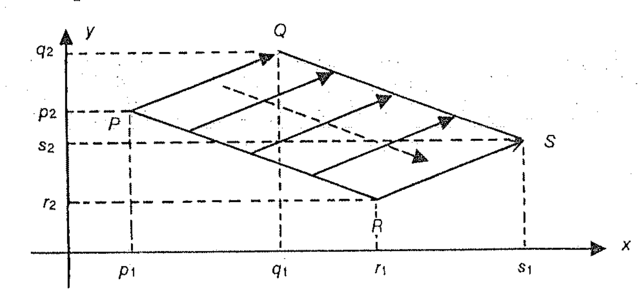
Recuérdese que se dijo que los vectores (x, y) de $\mathbf{R}^2$ podrían ser identificados geométricamente con segmentos dirigidos que parten del origen y llegan al punto (x, y). Por lo tanto, en principio, el segmento dirigido $\overrightarrow{PQ}$ no sería un vector de $\mathbf{R}^2$.

Sin embargo, esto no es así. Véase cómo proceder para identificar a $\overrightarrow{PQ}$ con un vector de $\mathbf{R}^2$.

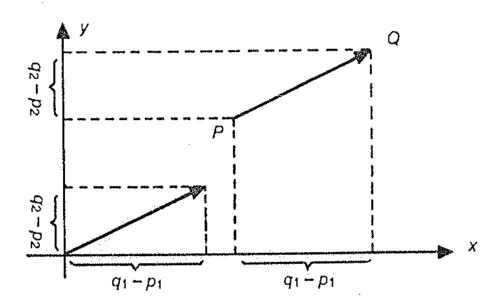
Sean $P = (p_1, p_2)$, $Q = (q_1, q_2)$, $R = (r_1, r_2)$, $S = (s_1, s_2)$ puntos del plano xy. Considérense los segmentos dirigidos $\overrightarrow{PQ}$ y $\overrightarrow{RS}$. Se dirá que estos segmentos dirigidos son *equivalentes* si

$$(q_1 - p_1,\ q_2 - p_2) = (s_1 - r_1,\ s_2 - r_2)$$

geométricamente, el hecho de que $\overrightarrow{PQ}$ sea equivalente a $\overrightarrow{RS}$ significa que por medio de un movimiento rígido (sin alterar el tamaño de $\overrightarrow{PQ}$) y paralelo a la dirección de $\overrightarrow{PQ}$, se puede hacer que este segmento coincida con $\overrightarrow{RS}$.



Es fácil ver que esta relación entre segmentos dirigidos del plano es una relación de equivalencia (es *reflexiva*, es *simétrica* y es *transitiva*). Por lo tanto, se puede identificar (como siendo los mismos) segmentos dirigidos que sean equivalentes. En particular, el segmento dirigido $\vec{PQ}$ es equivalente al *vector* $(q_1 - p_1, q_2 - p_2)$ (es decir, al segmento dirigido cuyo punto inicial es el origen y cuyo punto final es el punto $(q_1 - p_1, q_2 - p_2)$).



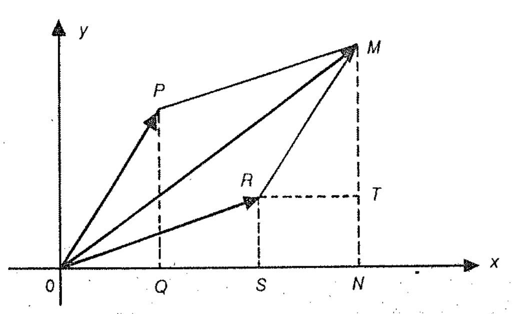
Se tiene entonces así una identificación completa entre vectores del espacio vectorial $\mathbf{R}^2$ y segmentos dirigidos del plano xy (a partir de este momento, también se usará la palabra vector para denotar los segmentos dirigidos del plano xy).

El espacio $\mathbf{R}^2$ es entonces un espacio "que se puede ver" (geométricamente).

Véase ahora el carácter geométrico de las operaciones definidas en este espacio vectorial.

Sean $v_1 = (x_1, y_1)$ y $v_2 = (x_2, y_2)$ dos vectores de $\mathbf{R}^2$. Se definió la suma de éstos como el vector $v_1 + v_2$ de $\mathbf{R}^2$ de coordenadas $(x_1 + y_1, x_2 + y_2)$. Se afirma que ésta no es más que la definición (abstracta) de la vieja conocida "regla del paralelogramo" que se usaba para sumar vectores en el plano.

En efecto, considérese la figura



Sean (a, b) las coordenadas del vector suma de $v_1 = \vec{OP}$ y $v_2 = \vec{OR}$ obtenido geométricamente con la ley del paralelogramo (es decir, si se completa el paralelogramo con v_1 y v_2 como lados, el vector suma $v_1 + v_2$ se obtiene uniendo el origen con el vértice opuesto del paralelogramo construido).

De la figura es claro que

$$|\overline{ON}| = |\overline{OS}| + |\overline{SN}|$$

Obsérvese que los triángulos OPQ y RMT son congruentes. Entonces $|\overline{OQ}| = |\overline{RT}| = |\overline{SN}|$, por lo que

$$|\overline{ON}| = |\overline{OS}| + |\overline{OQ}|$$

O sea

$$a = x_1 + x_2$$

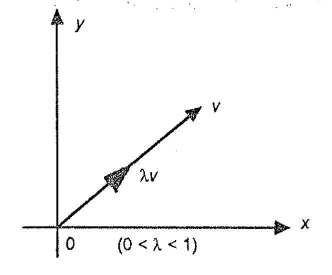
Un argumento similar conduce a

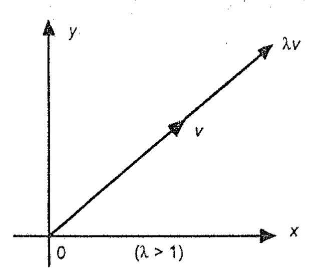
$$b = y_1 + y_2$$

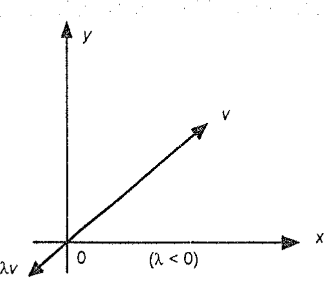
lo que prueba la afirmación.

Considérese ahora la multiplicación por escalares. Sea $v = (x, y)$ un vector de $\mathbf{R}^2$. Se ha definido la multiplicación de v por el escalar $\lambda \in \mathbf{R}$ como siendo el vector $\lambda v \in \mathbf{R}^2$ de coordenadas $(\lambda x, \lambda y)$. El efecto geométrico de esta operación consiste en, sin salir de la línea de acción del vector v:

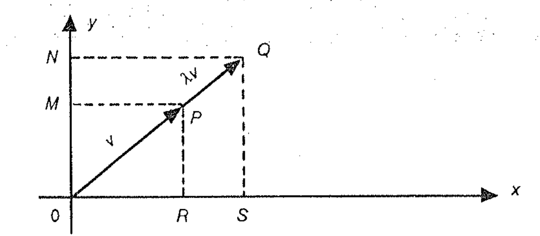
1) reducir su tamaño en un factor λ conservando su sentido $(0 < \lambda < 1)$;
2) aumentar su tamaño en un factor λ conservando su sentido $(\lambda > 1)$;
3) cambiar su sentido $(\lambda < 0)$.







En efecto, véase el caso particular $\lambda > 0$. Sean (a, b) las coordenadas del vector λv. El vector λv se obtiene geométricamente colocando sobre v (con su mismo sentido) el vector cuyo tamaño sea λ veces el tamaño de v. En la figura, localícese el punto Q sobre la dirección de $\vec{OP}$ de modo que $|\overline{OQ}| = \lambda |\overline{OP}|$



Los triángulos OPR y OQS son semejantes, de modo que

$$\frac{|\overline{OS}|}{|\overline{OR}|} = \frac{|\overline{SQ}|}{|\overline{RP}|} = \frac{|\overline{OQ}|}{|\overline{OP}|} = \lambda$$

o sea

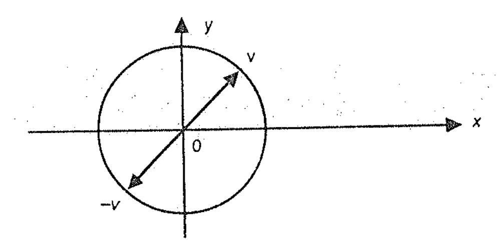
$$\frac{a}{x} = \frac{b}{y} = \lambda$$

esto es

$$a = \lambda x, \qquad b = \lambda y$$

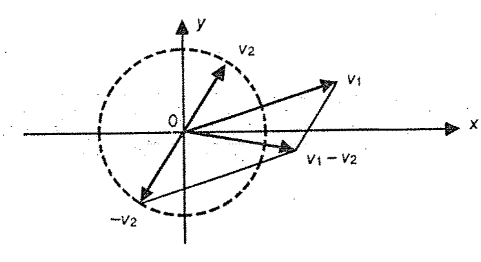
lo que prueba que el vector λv obtenido con la construcción geométrica anterior es el vector producto de v por el escalar λ, según la definición de esta operación en $\mathbf{R}^2$.

En particular, obsérvese que se tiene la siguiente relación geométrica entre el vector v y su inverso aditivo $(-v)$

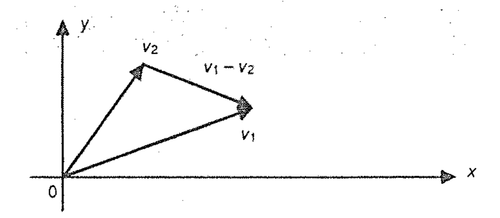


Dados los vectores $v_1 = (x_1, y_1)$, $v_2 = (x_2, y_2)$ en $\mathbf{R}^2$, se puede definir la *resta* $v_1 - v_2$ como $v_1 + (-v_2)$. Por ejemplo, si $v_1 = (7, 4)$ y $v_2 = (3, 1)$, se tiene $v_1 - v_2 = (7 - 3, 4 - 1) = (4, 3)$.

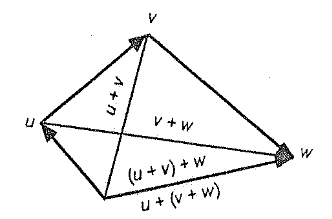
Geométricamente, el vector resta $v_1 - v_2$ se obtiene sumando (según como fue definida la resta) el vector v_1 con el inverso aditivo de v_2 según la regla del paralelogramo como se muestra en la figura



Obsérvese que un vector equivalente al vector $v_1 - v_2$ obtenido en la figura anterior es aquél cuyo punto inicial está en el punto final de v_2 y cuyo punto final está en el punto final de v_1



La propiedad asociativa de la suma de vectores en $\mathbf{R}^2$ también puede verse geométricamente. Ésta se ilustra en la siguiente figura



II. El espacio $\mathbf{R}^3$

El espacio $\mathbf{R}^3$ está constituido por ternas ordenadas de números reales, dígase (x, y, z) con $x, y, z \in \mathbf{R}$. O sea

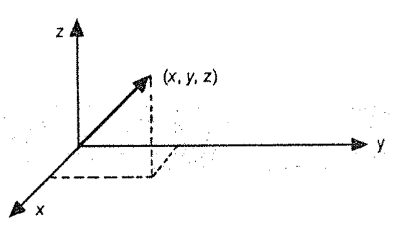
$$\mathbf{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbf{R}\}$$

Las operaciones de espacio vectorial se ven en este caso como

$$(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$$

$$\lambda(x, y, z) = (\lambda x, \lambda y, \lambda z), \lambda \in \mathbf{R}$$

Cada vector (x, y, z) de este espacio se puede identificar con el vector (geométrico) en el espacio tridimensional xyz cuyo punto inicial es el origen y cuyo punto final es el punto de coordenadas (x, y, z).



Todas las propiedades geométricas discutidas para el espacio vectorial $\mathbf{R}^2$ se tienen también en este espacio: la suma de vectores en $\mathbf{R}^3$ se hace de acuerdo a la regla del paralelogramo (en el plano en el que se encuentran los vectores sumandos), la multiplicación por escalares tiene el efecto geométrico de alterar el tamaño del

vector dejándolo con mismo sentido ($\lambda > 0$) o invirtiendo el sentido ($\lambda < 0$), etc. Se deja como ejercicio al lector que reproduzca la discusión geométrica que se hizo en el caso de $\mathbf{R}^2$, con el espacio vectorial $\mathbf{R}^3$.

En el caso general del espacio $\mathbf{R}^n$, se podría también establecer una geometría para los vectores que lo constituyen: un vector $(x_1, x_2, \ldots, x_n)$ de $\mathbf{R}^n$ se puede identificar con el vector (geométrico) cuyo punto inicial está en el origen (el cero del espacio $\mathbf{R}^n$) y cuyo punto final está en el punto $(x_1, x_2, \ldots, x_n)$ del espacio n-dimensional.

El problema que surge con este sentido geométrico de $\mathbf{R}^n$, es que no se puede *ver* (sólo "imaginar") a los vectores para $n \geq 4$.

NOTA: Se ha visto que $\mathbf{R}^n$ es un espacio vectorial real. Se consideró la geometría de los casos particulares $n = 2$ y $n = 3$. El caso particular $n = 1$ no fue considerado, pues, bajo la perspectiva de espacios vectoriales, resulta poco interesante el hecho de que $\mathbf{R}$ es un espacio vectorial sobre $\mathbf{R}$, con las operaciones operaciones usuales de suma y producto por reales. En realidad, es un hecho general que un campo K es un espacio vectorial sobre K. La razón por la cual no se hace hincapié en el caso particular $n = 1$ es, pues, que si bien $\mathbf{R}$ es un espacio vectorial sobre $\mathbf{R}$, éste es mucho más que eso: es un campo (cosa que no ocurre con $\mathbf{R}^n$, $n \geq 2$).

2.2. EL ESPACIO P_n

Sea P_n el conjunto de todos los polinomios con coeficientes reales de grado menor o igual que n. Es decir

$$P_n = \{a_0 + a_1x + \ldots + a_nx^n \mid a_0, a_1, \ldots, a_n \in \mathbf{R}\}^*$$

En este conjunto, se define la suma de polinomios "coeficiente a coeficiente", o sea, si $p_1 = a_0 + a_1x + \ldots + a_nx^n$ y $p_2 = b_0 + b_1x + \ldots + b_nx^n$ son dos polinomios de P_n, su suma es el polinomio $p_1 + p_2$ definido por

$$p_1 + p_2 = (a_0 + b_0) + (a_1 + b_1)x + \ldots + (a_n + b_n)x^n$$

Obsérvese que $p_1 + p_2 \in P_n$.

Para el polinomio $p = a_0 + a_1x + \ldots + a_nx^n \in P_n$ y el escalar $\lambda \in \mathbf{R}$, se define el producto de p por λ, como siendo el polinomio

$$\lambda p = \lambda a_0 + \lambda a_1 x + \ldots + \lambda a_n x^n$$

Nuevamente obsérvese que $\lambda p \in P_n$.

*La igualdad en este conjunto se define como: $p_1 = a_0 + a_1x + \ldots + a_nx^n$ es igual a $p_2 = b_0 + b_1x + \ldots + b_nx^n$, si, y sólo si $a_0 = b_0$, $a_1 = b_1, \ldots, a_n = b_n$.

Se puede verificar fácilmente que P_n con estas operaciones es un espacio vectorial: el cero de este espacio es el polinomio que tiene todos sus coeficientes iguales a cero, al que se denota simplemente por 0; esto es, $0 = 0 + 0x + \ldots + 0x^n$. Dado el polinomio $p = a_0 + a_1x + \ldots + a_nx^n \in P_n$, su inverso aditivo es el polinomio $-p = -a_0 - a_1x - \ldots - a_nx^n \in P_n$, etc.

2.3 EL ESPACIO $M_{m \times n}$

Sea $M_{m \times n}$ el conjunto de todas las matrices de orden $m \times n$. Es decir

$$M_{m\times n} = \left\{A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} \mid a_{ij} \in \mathbf{R},\ i = 1, \ldots, m, j = 1, \ldots, n\right\}$$

En el capítulo 1 se ha definido la suma de elementos de $M_{m \times n}$ así como la multiplicación de un elemento $M_{m \times n}$ por el escalar $\lambda \in \mathbf{R}$ (sección 3.1 del capítulo 1).

El teorema 3.1 del capítulo 1 afirma que el conjunto $M_{m \times n}$, con estas operaciones, es un espacio vectorial.

2.4. EL ESPACIO $F(I)$

Sea $F(I)$ el conjunto de todas las funciones reales definidas en el subconjunto I de la recta real. Es decir

$$F(I) = \{f: I \to \mathbf{R}\}^*$$

Sean f y g dos funciones de $F(I)$. Defínase la suma de f y g como la función $f + g: I \to \mathbf{R}$ tal que

$$(f + g)(x) = f(x) + g(x)$$

(el valor de la función $f + g$ en el punto $x \in I$, es la suma de los valores de las funciones f y g en ese punto).

Si $f \in F(I)$ y $\lambda \in \mathbf{R}$, defínase el producto de f por el escalar λ como la función $\lambda f: I \to \mathbf{R}$ tal que

$$(\lambda f)(x) = \lambda f(x)$$

(el valor de la función λf en el punto $x \in I$, es el producto de λ por el valor (el número real) de f en x).

Se verificará que el conjunto $F(I)$ con estas operaciones de suma y producto por escalares es un espacio vectorial. Sean $f, g, h \in F(I)$ y $\lambda, \mu \in \mathbf{R}$.

1) La suma es *conmutativa*.

*La igualdad entre elementos de este conjunto se define como: $f, g \in F(I), f = g$ si, y sólo si $f(x) = g(x)\ \forall x \in I$.

Sea $x \in I$. Entonces

$$(f+g)(x) = f(x) + g(x) = g(x) + f(x) = (g+f)(x)$$

como $(f+g)(x) = (g+f)(x)\ \forall\ x \in I$, concluimos que $f+g = g+f$ (es decir, la función $f+g$ es igual a la función $g+f$).

2) La suma es *asociativa*.

Sea $x \in I$. Entonces

$$\begin{aligned}(f+(g+h))(x) &= f(x) + (g+h)(x) = f(x) + (g(x) + h(x)) \\ &= (f(x) + g(x)) + h(x) = (f+g)(x) + h(x) \\ &= ((f+g)+h)(x)\end{aligned}$$

Nuevamente, como $x \in I$ fue arbitrario, se concluye que

$$f+(g+h) = (f+g)+h$$

3) Considérese la función $0: I \to \mathbf{R}$, definida por $0(x) = 0\ \forall\ x \in I$.
Obsérvese que para $x \in I$

$$(f+0)(x) = f(x) + 0(x) = f(x) + 0 = f(x)$$

Es decir, $f + 0 = f$. La función 0 es el cero de este espacio vectorial.

4) Dada $f \in F(I)$, defina la función $(-f): I \to \mathbf{R}$ como $(-f)(x) = -f(x)$. Entonces, para $x \in I$

$$(f+(-f))(x) = f(x) + (-f)(x) = f(x) - f(x) = 0$$

O sea, $f + (-f) = 0$, y entonces $(-f)$ es el inverso aditivo de f.

5) Para $x \in I$ se tiene

$$\begin{aligned}(\lambda(f+g))(x) &= \lambda(f+g)(x) = \lambda(f(x) + g(x)) = \lambda f(x) + \lambda g(x) \\ &= (\lambda f)(x) + (\lambda g)(x) = ((\lambda f) + (\lambda g))(x)\end{aligned}$$

Es decir $\lambda(f+g) = \lambda f + \lambda g$.

6) Si $x \in I$

$$((\lambda+\mu)f)(x) = (\lambda+\mu)f(x) = \lambda f(x) + \mu f(x) = (\lambda f)(x) + (\mu f)(x) = ((\lambda f) + (\mu f))(x)$$

O sea $(\lambda+\mu)f = \lambda f + \mu f$.

7) Si $x \in I$, $$\begin{aligned}((\lambda\mu)f)(x) &= (\lambda\mu)f(x) = \lambda(\mu f(x)) \\ &= \lambda((\mu f)(x)) = (\lambda(\mu f))(x)\end{aligned}$$

en donde $(\lambda\mu)f = \lambda(\mu f)$.

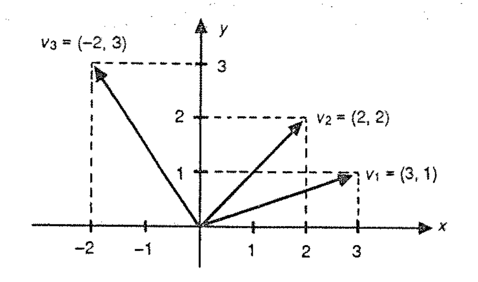
8) Para $x \in I$, $(1 \cdot f)(x) = 1 \cdot f(x) = f(x)$, es decir, $1 \cdot f = f$.

Esto demuestra, entonces que $F(I)$ es un espacio vectorial.

Los ejemplos de espacios vectoriales aquí dados muestran solamente algunos de los muchos espacios con los que se trabajará en el resto del libro. Estos nuevos espacios podrán obtenerse de los ejemplos aquí estudiados considerando "porciones" de ellos. En la siguiente sección se estudiará cómo han de ser estas "porciones" para que ellas mismas sean también espacios vectoriales.

EJERCICIOS (SECCIÓN 2, CAPÍTULO 3)

1. Escriba explícitamente
   a) El vector cero de $\mathbf{R}^4$.
   b) El inverso aditivo del vector $(1, -2, 0, 3, 8)$ en $\mathbf{R}^5$.
   c) El vector $(1, 3, 2) + ((1, 4, 3) + (2, -1, -3))$ en $\mathbf{R}^3$.
   d) El vector cero de $\mathbf{R}^7$.
   e) El vector $5(8, -1, 3, 2)$ en $\mathbf{R}^4$.
   f) El inverso aditivo del vector $(1, 1, 1, 3, 8, -1, 7)$ en $\mathbf{R}^7$ multiplicado por el escalar 4.
   g) El vector $(1, 3, 8, -1, 4) + (4, 2, -1, 7, 0)$ en $\mathbf{R}^5$ multiplicado por el escalar -3.
   h) El inverso aditivo del vector $2(1, -1, 8, 5, 10, 4, 0, 1)$ en $\mathbf{R}^8$.
   i) El vector $(1, 1, 3, 27, 45)$ de $\mathbf{R}^5$ multiplicado por el escalar 0.
   j) El inverso aditivo del vector $(1, 1, 1, 1, 1, 1)$ de $\mathbf{R}^6$ multiplicado por el escalar -1.
   k) La propiedad conmutativa de la suma de vectores en $\mathbf{R}^4$.
   l) El vector $2(1, 1, -1, 3) + 4(8, 1, 7, -1)$ de $\mathbf{R}^4$.
   m) El vector $(5, 7, 9, 10, 45, 86, -70)$ de $\mathbf{R}^7$ multiplicado por el escalar 1.
   n) El vector $5(1, 3, -1, 4, 8)$ de $\mathbf{R}^5$ multiplicado por el escalar -2.
   o) El vector de $\mathbf{R}^5$ multiplicado por el escalar 3 da por resultado el vector $(7, 1, -1, 0, 4)$.
2. Demuestre que la relación entre segmentos dirigidos en el plano xy establecida en la página 187 es una relación de equivalencia.
3. Considere los segmentos dirigidos en el plano xy mostrados en la siguiente figura:



A continuación se dan el punto inicial (*PI*) y el punto final (*PF*) de varios segmentos dirigidos en el plano *xy*. Diga en cada caso si el segmento dirigido es equivalente a v_1, a v_2, a v_3 o a ninguno de ellos.

a) PI = (1, 6)	PF = (3, 8)	e) PI = (-10, 4)	PF = (-8, 6)
b) PI = (2, 1)	PF = (3, 4)	f) PI = (3, 9)	PF = (6, 10)
c) PI = (12, -4)	PF = (10, -1)	g) PI = (6, 10)	PF = (4, 13)
d) PI = (4, -2)	PF = (7, -1)	h) PI = (1, 3)	PF = (0, 4)

4. Demuestre geométricamente (con la regla del paralelogramo) que la suma del vector $(a, 0)$ con el vector $(0, b)$, en donde a y b son números positivos, es el vector (a, b).

5. Dado el vector $v = (1, 2)$ en el plano *xy*, haga una gráfica mostrando el vector

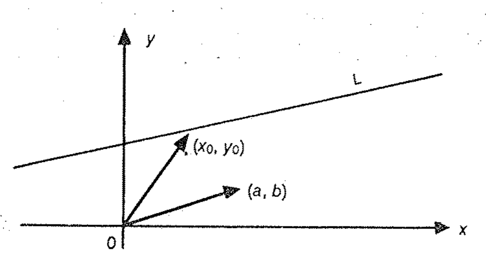
a) $2v$	c) $-v$	e) $-3v$
b) $1.5v$	d) $-2v$	f) $3v$

6. Dado el vector no nulo $v = (a, b)$ y el vector $v_0 = (x_0, y_0)$, demuestre que la recta que pasa por v_0 y es paralela a la recta en la que se encuentra v puede ser descrita de la siguiente manera

$$L = \{(x, y) \mid (x, y) = (x_0, y_0) + t(a, b),\ t \in \mathbf{R}\}$$

de donde se puede deducir que la ecuación de L puede escribirse como

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \end{cases} \quad t \in \mathbf{R}$$



Describa geométricamente las siguientes rectas en el plano:

a) $\begin{cases} x = 2 + t \\ y = -1 + t \end{cases} \quad t \in \mathbf{R}$

b) $\begin{cases} x = 3t \\ y = -2 - t \end{cases} \quad t \in \mathbf{R}$

c) $\begin{cases} x = -3 + s \\ y = 3 - s \end{cases} \quad s \in \mathbf{R}$

d) $\begin{cases} x = t \\ y = t \end{cases} \quad t \in \mathbf{R}$

e) $\begin{cases} x = 4t \\ y = 4t \end{cases} \quad t \in \mathbf{R}$

f) $\begin{cases} x = 3t + 1 \\ y = 3t + 1 \end{cases} \quad t \in \mathbf{R}$

7. Considere las rectas L_1 y L_2 dadas por

$$L_1: \begin{cases} x = x_0 + at \\ y = y_0 + bt \end{cases} \quad t \in \mathbf{R} \qquad L_2: \begin{cases} x = x_0' + a't' \\ y = y_0' + b't' \end{cases} \quad t' \in \mathbf{R}$$

a) Compruebe que L_1 y L_2 se cortan en un solo punto si, y sólo si $a'b - ab' \neq 0$.

b) Demuestre que si $a'b - ab' \neq 0$, el punto en donde L_1 y L_2 se cortan es

$$x = x_0 + a\,\frac{a'(y_0' - y_0) - b'(x_0' - x_0)}{a'b - ab'}$$

$$y = y_0 + b\,\frac{a'(y_0' - y_0) - b'(x_0' - x_0)}{a'b - ab'}$$

8. Dado el vector no nulo $v = (a, b, c)$ en $\mathbf{R}^3$ y el vector $v_0 = (x_0, y_0, z_0)$, demuestre que la recta que pasa por v_0 y es paralela a la recta en la que se encuentra el vector v puede ser descrita como

$$L = \{(x, y, z) \mid (x, y, z) = (x_0, y_0, z_0) + t(a, b, c) \quad t \in \mathbf{R}\}$$

Es decir, que la ecuación de L, puede escribirse como

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = z_0 + ct \end{cases} \quad t \in \mathbf{R}$$

9. Al usar argumentos geométricos, compruebe que la ecuación de la recta $\mathbf{R}^3$ que pasa por los puntos (x_0, y_0, z_0) y (x_1, y_1, z_1) se puede escribir como

$$\begin{cases} x = x_0 + t(x_1 - x_0) \\ y = y_0 + t(y_1 - y_0) \\ z = z_0 + t(z_1 - z_0) \end{cases} \quad t \in \mathbf{R}$$

10. Escriba explícitamente

a) El vector cero de P_3.
b) El inverso aditivo del vector $1 + x + x^2$ de P_2.
c) La suma del vector x^5 y el vector $2x^5 + x + 3$ en P_5.
d) El producto del vector $x^3 + 8x^2 + 3x + 1$ por el escalar -3.
e) El vector $(-4 + 8x + x^2) + (-1 - 10x - 2x^2)$ en P_2 por el escalar -1.
f) El inverso aditivo del vector $3(2 - x) - 10(4 + x)$ en P_1.

11. Escriba explícitamente
a) El vector cero de $M_{1\times 4}$.
b) El vector cero de $M_{3\times 5}$
c) El inverso aditivo del vector $\begin{bmatrix} -1 & 1 \\ 2 & 3 \end{bmatrix}$ en $M_{2\times 2}$
d) El vector $2\begin{bmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \end{bmatrix} + 4\begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}$ en $M_{2\times 3}$
e) El inverso ditivo del vector [1 3] + 8[2 -1] en $M_{1\times 2}$ multiplicado por el escalar -1.
f) El vector de $M_{3\times 2}$ que sumado con el vector $3\begin{bmatrix} 1 & -2 \\ 3 & 8 \\ 5 & -4 \end{bmatrix}$ dé por resultado el vector $M_{3\times 2}$ cuyos elementos son todos nulos.

12. a) Dibuje la gráfica del vector cero del espacio vectorial $F([a, b])$, en donde $a < b$.
b) Dibuje la gráfica del inverso aditivo del vector $f(x) = 1$ en $F(\mathbf{R})$.
c) Dibuje la gráfica del inverso aditivo del vector $f(x) = \operatorname{sen} x$ en $F(\mathbf{R})$.
d) Demuestre que los únicos valores de c_1 y c_2 tales que el vector $c_1e^x + c_2xe^x$ es igual al vector cero (en $F(\mathbf{R})$) son $c_1 = c_2 = 0$.
e) Pruebe que existen valores no nulos para c_1, c_2 y c_3 tales que el vector $c_1(x + 1) + c_2(3x + 2) + c_3(-x + 4)$ es igual al vector cero de $F(\mathbf{R})$.

3. SUBESPACIOS

Sea V un *espacio vectorial* y considérese un subconjunto W de V (es decir, el conjunto W es tal que todo elemento de W es elemento de V). W hereda de manera natural, las operaciones de espacio vectorial que estaban definidas en V. Se tiene así, un conjunto W en donde existen operaciones de suma y producto por escalares. ¿Es W un espacio vectorial?

Véanse un par de ejemplos:

EJEMPLO 1 Sea V el espacio vectorial $\mathbf{R}^2$. Sea W el subconjunto de $\mathbf{R}^2$ formado por todos aquellos vectores (x, y) que tiene su segunda coordenada positiva (geométricamente W corresponde al semiplano superior del plano xy). Es decir,

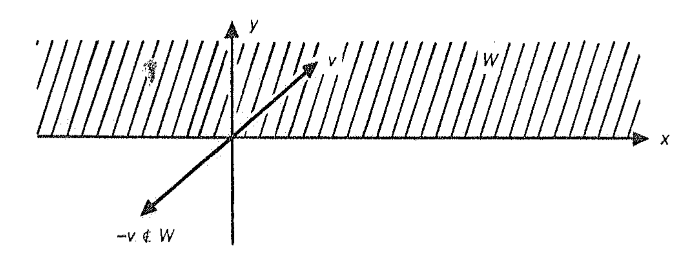
$$W = \{(x, y) \mid x, y \in \mathbf{R}, y > 0\}$$

Las operaciones entre los elementos de W son las que estaban definidas en $\mathbf{R}^2$, es decir

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$
$$\lambda(x, y) = (\lambda x, \lambda y), \qquad \lambda \in \mathbf{R}$$

Uno se pregunta si W es un espacio vectorial.

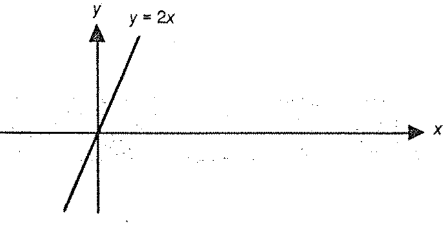
De hecho, existen muchas razones por las que la respuesta a esta pregunta es NO. En principio, la operación de producto por escalares no es cerrada, pues si se toma el escalar $\lambda < 0$, el vector $\lambda(x, y) = (\lambda x, \lambda y)$ no pertenece a W, pues siendo $y > 0$, se tendría $\lambda y < 0$. Consecuencia de este hecho es que dado un vector (x, y) en W, éste no tiene (en W) inverso aditivo. Geométricamente esto se ve como



EJEMPLO 2 Sin embargo, si W es el siguiente subconjunto de $\mathbf{R}^2$

$$W = \{(x, y) \mid x, y \in \mathbf{R}, y = 2x\}$$

[o sea, W es el conjunto de parejas ordenadas del tipo $(x, 2x)$, $x \in \mathbf{R}$], se puede ver que en este caso W sí es un espacio vectorial con las operaciones usuales de $\mathbf{R}^2$. En principio observe que si $u = (x_1, y_1)$ y $v = (x_2, y_2)$ son dos vectores de W (o sea $y_1 = 2x_1$ y $y_2 = 2x_2$) entonces $u + v = (x_1 + x_2, y_1 + y_2)$ también pertenece a W pues $y_1 + y_2 = 2x_1 + 2x_2 = 2(x_1 + x_2)$. Similarmente, si $(x, y) \in W$ y $\lambda \in \mathbf{R}$, se tiene $\lambda(x, y) = (\lambda x, \lambda y) \in W$ pues $\lambda y = \lambda(2x) = 2(\lambda x)$. La verificación para W de los 8 axiomas de la definición de espacio vectorial se deja como un sencillo ejercicio para el lector. Geométricamente W está representado por la recta $y = 2x$ en el plano xy



DEFINICIÓN 3.1 Sea V un espacio vectorial y W un subconjunto de V. Se dice que W es un subespacio de V si W es por sí mismo un espacio vectorial con las operaciones de suma y producto por escalares que estaban definidas en V.

Así pues, el conjunto $\{(x, y) \mid x, y \in \mathbf{R}, y > 0\}$ no es un subespacio de $\mathbf{R}_2$, mientras que el conjunto $\{(x, y) \mid x, y \in \mathbf{R}, y = 2x\}$ sí lo es.

Obsérvese que todo espacio vectorial tiene al menos dos subespacios, a saber, el subespacio $\{0\}$ y V mismo. A éstos se les llamará subespacios triviales de V.

En principio, para poder saber si un subconjunto W del espacio vectorial V, es subespacio de éste o no, habría que verificar (la cerradura de las operaciones en W y) los 8 axiomas que definen a un espacio vectorial en el subconjunto W.

Sin embargo, el teorema siguiente dice que tal verificación requiere de mucho menos trabajo.

TEOREMA 3.1

> Sea V un espacio vectorial y sea W un subconjunto no vacío de V. Entonces W es un subespacio de V si, y sólo si se cumplen las dos condiciones siguientes:
>
> a) $u, v \in W \Rightarrow u + v \in W$
>
> b) $v \in W, \lambda \in \mathbf{R} \Rightarrow \lambda v \in W$

DEMOSTRACIÓN La necesidad de las condiciones a) y b) es obvia, pues siendo W un subespacio de V, éste es un espacio vectorial por sí mismo. En particular, deben cumplirse las condiciones a) y b). Se verá que estas condiciones también son suficientes para garantizar que W es subespacio de V. Se tendrá que verificar que W es un espacio vectorial. Obsérvese que las condiciones a) y b) dicen que las operaciones de suma y producto por escalares son cerradas en W (esto es, se tiene $+ : W \times W \to W$, y $\cdot : R \times W \to W$).

Algunos de los axiomas de la definición de espacio vectorial se cumplirán automáticamente en W por el simple hecho de ser W un subconjunto de V. Por ejemplo, el axioma 1, dice que para todo par de vectores $u, v \in W$ debe cumplirse $u + v = v + u$. Pero $u, v \in V$ y en V se tiene la conmutatividad de la suma, por tanto, en W se cumple también este axioma. Con este mismo argumento se puede ver que los axiomas 2), 5), 6), 7) y 8) se verifican en W. Véase que W satisface también el axioma 3): se sabe que existe el elemento $0 \in V$ y se quiere ver que éste pertenece de hecho a W. Por la propiedad b), tomando $\lambda = 0$ y v cualquier vector de W, se tiene $\lambda \cdot v = 0 \cdot v \in W$. Pero $0 \cdot v = 0$ (teorema 1.1 (1)), y entonces $0 \in W$. Por último, se tiene también el axioma 4) pues si $v \in W$, tomando $\lambda = -1$ en la condición b) se concluye $(-1)v \in W$. Pero $(-1)v = -v$ (teorema 1.2 (2)), o sea, $-v \in W$.

Q.E.D.

Con la ayuda de este teorema será, entonces, una tarea muy simple verificar si un subconjunto de un espacio vectorial es subespacio de él.

Véase un ejemplo interesante: se trata de ver cómo son los subespacios de $\mathbf{R}^2$.

EJEMPLO 3 Considere el conjunto L de $\mathbf{R}^2$ dado por

$$L = \{(x, y) \in \mathbf{R}^2 \mid ax + by = 0\}$$

en donde a y b son dos números reales fijos no ambos cero.

Se afirma que W es un subespacio de $\mathbf{R}^2$.

En efecto, sean $u = (x_1, y_1)$ y $v = (x_2, y_2)$ dos vectores de L. Se tendrá que verificar que $u + v \in L$ y que $\lambda u \in L$ para $\lambda \in \mathbf{R}$. Observe que $u + v = (x_1 + x_2, y_1 + y_2)$ y

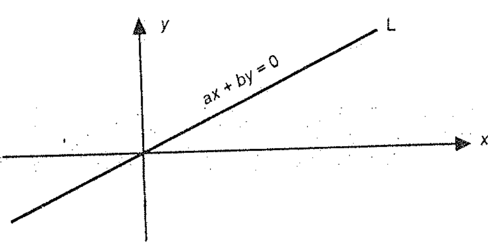
$$a(x_1 + x_2) + b(y_1 + y_2) = \underbrace{(ax_1 + by_1)}_{\substack{=0 \\ \text{pues } (x_1, y_1) \in L}} + \underbrace{(ax_2 + by_2)}_{\substack{=0 \\ \text{pues } (x_2, y_2) \in L}} = 0$$

lo que dice entonces que $u + v \in$ L. Similarmente $\lambda u = (\lambda x_1, \lambda y_1)$ y

$$a(\lambda x_1) + b(\lambda y_1) = \lambda(ax_1 + by_1) = \lambda 0 = 0$$

lo que prueba que $\lambda u \in L$.

Geométricamente el subespacio L es una recta en el plano xy que pasa por el origen



Obsérvese que L es un subespacio no trivial de $\mathbf{R}^2$ ($L \neq \{0\}, \mathbf{R}^2$). El siguiente teorema dice que $\mathbf{R}^2$ no tiene más subespacios de este tipo; es decir, todo subespacio no trivial de $\mathbf{R}^2$ es del tipo de L descrito anteriormente.

TEOREMA 3.2

> Sea W un subespacio no trivial de $\mathbf{R}^2$. Entonces W es de la forma
>
> $$W = \{(x, y) \mid ax + by = 0\}$$
>
> donde a y b son números reales no ambos nulos.

DEMOSTRACIÓN Sea (x_0, y_0) un vector no nulo de W. Dígase entonces que $x_0 \neq 0$.

Considérese el subconjunto de $\mathbf{R}^2$ dado por

$$L = \{(x, y) \mid y_0 x - x_0 y = 0\}$$

Se ha probado ya que L es un subespacio de $\mathbf{R}^2$. Se afirma que de hecho $L = W$.

Primeramente véase que $L \subset W$. Sea $(x_1, y_1) \in L$. Entonces $y_0 x_1 - x_0 y_1 = 0$, o bien, $y_1 = y_0 x_1 / x_0$. Sea $\lambda = x_1 / x_0 \in \mathbf{R}$.

Obsérvese que

$$(x_1, y_1) = \frac{x_1}{x_0}(x_0, y_0) = \lambda(x_0, y_0)$$

Pero $(x_0, y_0) \in W$ y W es un subespacio de $\mathbf{R}^2$. Entonces $(x_1, y_1) \in W$ lo que prueba $L \subset W$.

Recíprocamente, suponga por contradicción que $W \not\subset L$.

Existiría entonces $(x_2, y_2) \in W$ tal que $(x_2, y_2) \notin L$. Es decir, $(x_2, y_2) \in W$ y $y_0x_2 - x_0y_2 \neq 0$. Sea (x, y) un vector *cualquiera* de $\mathbf{R}^2$. Considérese el sistema de dos ecuaciones con dos incógnitas λ y μ siguiente

$$y_0\lambda + y_2\mu = y$$
$$x_0\lambda + x_2\mu = x$$

El determinante de la matriz del sistema es $y_0x_2 - x_0y_2$, el cual se está suponiendo distinto de cero. Entonces, independientemente de los valores de x y y (términos independientes) existe una única solución al sistema, $\tilde{\lambda}$ y $\tilde{\mu}$ dígase. Es decir, existen $\tilde{\lambda}$ y $\tilde{\mu}$ para todo x y todo y, tales que

$$x = x_0\tilde{\lambda} + x_2\tilde{\mu}$$
$$y = y_0\tilde{\lambda} + y_2\tilde{\mu}$$

Pero entonces

$$(x, y) = (x_0\tilde{\lambda} + x_2\tilde{\mu},\ y_0\tilde{\lambda} + y_2\tilde{\mu}) = \tilde{\lambda}(x_0, y_0) + \tilde{\mu}(x_2, y_2)$$

Tanto (x_0, y_0) como (x_2, y_2) son vectores de W. Siendo éste un subespacio de $\mathbf{R}^2$, combinando las condiciones a) y b) del teorema 3.1, vemos que $(x, y) \in W$. Pero (x, y) era un vector *arbitrario* de $\mathbf{R}^2$. Por lo tanto, $W = \mathbf{R}^2$, lo que contradice la hipótesis de que W es un subespacio no trivial de $\mathbf{R}^2$.

Entonces $W = L$. Tómese $a = y_0$, $b = -x_0$ y obtenga así la conclusión del teorema.

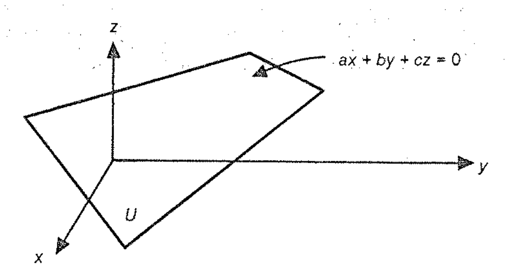
Q.E.D.

Continúese viendo algunos otros ejemplos de subespacios.

EJEMPLO 4 Considérese el subconjunto U de $\mathbf{R}^3$ dado por

$$U = \{(x, y, z) \mid ax + by + cz = 0\}$$

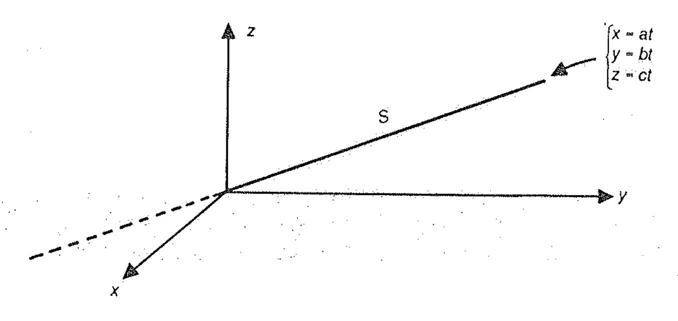
en donde a, b y c son tres números reales fijos no todos iguales a cero. Se verifica fácilmente que U satisface las condiciones a) y b) del teorema 3.1. Éste es pues un subespacio (no trivial) de $\mathbf{R}^3$. Obsérvese que geométricamente U representa un plano en el espacio tridimensional que pasa por el origen.



EJEMPLO 5 Se puede ver también que el subconjunto de $\mathbf{R}^3$

$$S = \{(x, y, z) \mid x = at, y = bt, z = ct, t \in \mathbf{R}\}$$

es un subespacio de $\mathbf{R}^3$. Geométricamente, S representa una recta que pasa por el origen (ver ejercicio 8 de la sección 2).



Se verá después (en la sección 5) que los únicos subespacios no triviales de $\mathbf{R}^3$ son de la forma de U o de S en los ejemplos anteriores.

Considérese ahora el espacio vectorial $F(I)$ introducido en la sección anterior.

Este espacio tiene muchos subespacios importantes en matemáticas. Véase uno de ellos.

Sea $C(I)$ el subconjunto de $F(I)$ formado por las funciones reales continuas definidas en I, es decir,

$$C(I) = \{f \in F(I) \mid f \text{ es continua}\}$$

Se afirma que $C(I)$ es un subespacio de $F(I)$.

La demostración de este hecho se basa en dos resultados que son estudiados en los cursos de cálculo, a saber: 1) la suma de funciones continuas es una función continua y 2) si se multiplica una función continua por el escalar μ, se obtiene una función continua. 1) y 2) son precisamente las condiciones a) y b) del teorema 3.1, que garantizan que $C(I)$ es un subespacio de $F(I)$. Se llamará a $C(I)$ el *espacio de funciones continuas*.

Algunos otros subespacios de $F(I)$ son considerados en los ejercicios al final de esta sección.

Por último, véase el siguiente ejemplo:

EJEMPLO 6 *Un ejemplo importante:*
(El espacio de soluciones de un sistema homogéneo de ecuaciones lineales.)

Considere el sistema homogéneo de m ecuaciones lineales con n incógnitas $AX = 0$ (A es entonces una matriz $m \times n$).

Identifique a la matriz X (de orden $n \times 1$) con el vector de $\mathbf{R}^n$ cuyas coordenadas son los elementos de esta matriz (en el mismo orden), es decir

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \xleftrightarrow{\text{identificación}} X = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$$

Se puede entonces decir que "el vector $X = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$ es una solución del sistema $AX = 0$", refiriéndose a que la matriz correspondiente, según la identificación convenida anteriormente, satisface al sistema.

Considere, entonces, el siguiente subconjunto de $\mathbf{R}^n$

$$S = \{(x_1, x_2, \ldots, x_n) \mid (x_1, x_2, \ldots, x_n) \text{ es solución de } AX = 0\}$$

Se afirma que S es un subespacio de $\mathbf{R}^n$.

En efecto, suponga que X_1 y X_2 son soluciones del sistema $AX = 0$. Entonces $A(X_1 + X_2) = AX_1 + AX_2 = 0 + 0 = 0$, lo que dice que $X_1 + X_2$ es también solución. Asimismo, si μ es un escalar, se tiene $A(\mu X_1) = \mu AX_1 = \mu 0 = 0$, de donde se ve que μX_1 es también solución. Entonces S es un subespacio de $\mathbf{R}^n$ que se llamará *espacio de soluciones* (o espacio solución) del sistema $AX = 0$.

Para ilustrar este último ejemplo, resuélvase el sistema homogéneo de ecuaciones lineales

$$x_1 + x_2 - 5x_3 = 0$$
$$x_1 + 2x_2 - 8x_3 = 0$$
$$4x_1 + 7x_2 - 29x_3 = 0$$

Al aplicar el método de eliminación Gaussiana se obtiene

$$\begin{bmatrix} 1 & 1 & -5 \\ 1 & 2 & -8 \\ 4 & 7 & -29 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}$$

de modo que el sistema original es equivalente a

$$x_1 - 2x_3 = 0$$
$$x_2 - 3x_3 = 0$$

Al realizar $x_3 = t$, se obtiene $x_1 = 2t$, $x_2 = 3t$, de modo que el conjunto solución del sistema, visto como un subconjunto de $\mathbf{R}^3$ es:

$$\{(x_1, x_2, x_3) \mid x_1 = 2t, x_2 = 3t, x_3 = t, \quad t \in \mathbf{R}\}$$

el cual es, como se había visto, un subespacio de $\mathbf{R}^3$ (una recta que pasa por el origen).

Por último, considere el sistema no homogéneo

$$x_1 + x_2 - 5x_3 = -10$$
$$x_1 + 2x_2 - 8x_3 = -17$$
$$4x_1 + 7x_2 - 29x_3 = -61$$

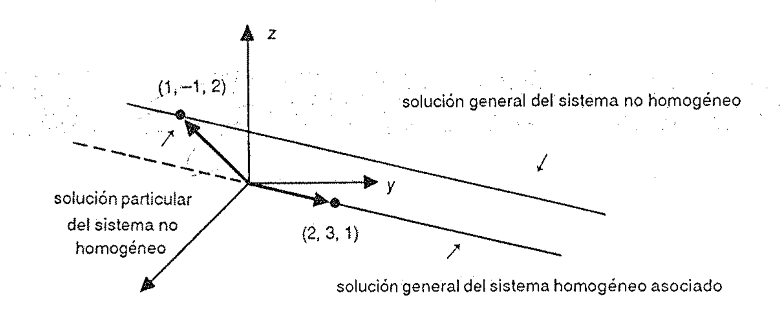
Según el teorema 3.3 de la sección 3 del capítulo 1, la solución general de este sistema está constituida por la suma de una solución particular de él y la solución general del sistema homogéneo asociado, el cual es precisamente el que se estudia en el ejemplo 6. Al observar que $x_1 = 1$, $x_2 = -1$ y $x_3 = 2$ es una solución particular del sistema, se ve que los vectores $(x_1, x_2, x_3) \in \mathbf{R}^3$ que constituyen su solución general son de la forma

$$(x_1, x_2, x_3) = \underbrace{(1, -1, 2)}_{\substack{\textbf{solución}\\\textbf{particular}}} + \underbrace{(2t, 3t, t)}_{\substack{\textbf{solución general}\\\textbf{del sistema}\\\textbf{homogéneo asociado}}}$$

Es decir que el conjunto solución de este sistema puede ser descrito como

$$\{(x_1, x_2, x_3) \mid (x_1, x_2, x_3) = (1, -1, 2) + t(2, 3, 1)\}$$

Según el ejercicio 8 de la sección anterior, este subconjunto de $\mathbf{R}^3$ representa geométricamente una recta que pasa por $(1, -1, 2)$ y es paralela a la recta en la que se encuentra el vector $(2, 3, 1)$. En otras palabras, la solución del sistema se representa geométricamente por una recta que es paralela a la recta en donde se encuentran las soluciones del sistema homogéneo asociado, desplazada por el "vector solución particular" del mismo.



3.1. SUBESPACIOS GENERADOS POR UN CONJUNTO DE VECTORES

En esta subsección se pretende discutir el siguiente problema, el cual jugará un papel muy importante dentro de la teoría que se desarrollará en las siguientes secciones: sea V un espacio vectorial y sea S un subconjunto (no vacío) de V. Uno se pregunta si existe algún subespacio de V que contenga a S.

Primeramente obsérvese que si S es un subespacio de V, la respuesta a nuestra pregunta anterior es obvia: S es un subespacio de V que contiene a S.

Aún más, aunque S no sea subespacio de V, el problema tiene una solución trivial: V es un subespacio de V que contiene a S.

Reestructúrese entonces la pregunta imponiendo una condición extra sobre el subespacio procurado: uno se pregunta si existe algún subespacio de V que contiene a S, que sea "el más pequeño de todos". Es decir, si se considera el conjunto de todos los subespacios de V que contienen a S (que no es vacío, pues V está en él), uno se pregunta si existe algún elemento minimal (en el sentido de contención de conjuntos) en él.

Para poder dar respuesta a esta pregunta, se tiene que introducir el importante concepto de "combinación lineal".

DEFINICIÓN 3.2 Sea V un espacio vectorial. Sean $v_1, v_2, \ldots, v_n$ vectores de V. Se dice que el vector $v \in V$ es una combinación lineal de los vectores $v_1, v_2, \ldots, v_n$, si existen escalares $c_1, c_2, \ldots, c_n$ tales que

$$v = c_1v_1 + c_2v_2 + \ldots + c_nv_n$$

EJEMPLO 7 Por ejemplo, el vector (14, 12, 2) en $\mathbf{R}^3$ es combinación lineal de los vectores (1, 2, -1) y (3, 2, 1) puesto que

$$(14, 12, 2) = 2(1, 2, -1) + 4(3, 2, 1)$$

En forma similar, el vector v de $M_{2\times 2}$ dado por

$$v = \begin{bmatrix} -3 & -4 \\ 2 & -5 \end{bmatrix}$$

es combinación lineal de los vectores

$$v_1 = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \quad \text{y} \quad v_2 = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$$

puesto que se puede escribir $v = 3v_1 - 2v_2$, como se puede comprobar directamente.

Regrese ahora al problema inicial. Se va a considerar el conjunto de todas las posibles combinaciones lineales de elementos de S. Denote por $\mathscr{L}(S)$ a este conjunto. Entonces

$$\mathscr{L}(S) = \{c_1v_1 + \ldots + c_kv_k \mid c_i \in \mathbf{R}, v_i \in S, i = 1, \ldots, k, \quad k \in \mathbf{N}\}$$

Note que cada elemento en $\mathscr{L}(S)$ es una combinación lineal de un número *finito* de elementos de S. Si S es un conjunto finito, diga que $S = \{v_1, \ldots, v_n\}$ entonces

$$\mathscr{L}(S) = \{c_1v_1 + \ldots + c_nv_n \mid c_i \in \mathbf{R}, i = 1, \ldots, n\}$$

Se puede demostrar que $\mathscr{L}(S)$ es un subespacio de V y que además, es el subespacio que resuelve el problema planteado inicialmente. El caso que interesa considerar en este libro es cuando S es un conjunto finito de vectores de V. Éste es el caso expuesto en el siguiente teorema:

TEOREMA 3.3 Sea V un espacio vectorial, y sean $v_1, v_2, \ldots, v_n$ n vectores de V. Entonces el conjunto

$$\mathscr{L}(v_1, v_2, \ldots, v_n) = \{c_1v_1 + \ldots + c_nv_n \mid c_i \in \mathbf{R}, \quad i = 1, \ldots, n\}$$

formado por todas las combinaciones lineales de $v_1, \ldots, v_n$ es un subespacio de V y además, es el menor de todos los subespacios de V que contienen a los vectores $v_1, v_2, \ldots, v_n$.

DEMOSTRACIÓN Se demostrará primeramente que $\mathscr{L}(v_1, \ldots, v_n)$ es un subespacio de V. Para esto se tendrá que verificar que $\mathscr{L}(v_1, \ldots, v_n)$ satisface las condiciones a) y b) del teorema 3.1. Sean $x, y \in \mathscr{L}(v_1, \ldots, v_n)$. Entonces existen $c_1, \ldots, c_n, d_1, \ldots, d_n \in \mathbf{R}$ tales que

$$x = c_1v_1 + \ldots + c_nv_n$$

$$y = d_1v_1 + \ldots + d_nv_n$$

Fórmese la suma $x + y$. Ésta es

$$x + y = (c_1 + d_1)v_1 + \ldots + (c_n + d_n)\, v_n$$

lo que demuestra que $x + y \in \mathscr{L}(v_1, \ldots, v_n)$.

En forma similar, si $\lambda \in \mathbf{R}$ se tiene

$$\lambda x = \lambda(c_1v_1 + \ldots + c_nv_n) = (\lambda c_1)v_1 + \ldots + (\lambda c_n)v_n$$

lo que prueba que $\lambda x \in \mathscr{L}(v_1, \ldots, v_n)$. Entonces, $\mathscr{L}(v_1, \ldots, v_n)$ es un subespacio de V.

Ciertamente $\mathscr{L}(v_1, \ldots, v_n)$ contiene a los vectores $v_1, \ldots, v_n$, pues

$$v_i = 0 \cdot v_1 + \ldots + 0 \cdot v_{i-1} + 1 \cdot v_i + 0 \cdot v_{i+1} + \ldots + 0 \cdot v_n \in \mathscr{L}(v_1, \ldots v_n)$$

$$i = 1, \ldots, n$$

Sea ahora W un subespacio de V que contiene a $v_1, \ldots, v_n$, y sea $x = c_1v_1 + \ldots + c_nv_n \in \mathscr{L}(v_1, \ldots, v_n)$. Por la condición de ser W un subespacio de V (aplicando las condiciones a) y b) del teorema 3.1 con los vectores $v_1, \ldots, v_n \in W$ y los escala-

res $c_1, \ldots, c_n$) se ve que $x \in W$. Es decir, $\mathcal{L}(v_1, \ldots, v_n) \subseteq W$. Esto demuestra que $\mathcal{L}(v_1, \ldots, v_n)$ es el menor de los subespacios de V que contiene a los vectores $v_1, \ldots, v_n$.

Q.E.D.

NOTA: La demostración del caso general: $\mathcal{L}(S)$ es el menor de los subespacios de V que contienen al subconjunto (no necesariamente finito) S; se puede obtener realizando algunas pequeñas modificaciones (obvias) en la demostración anterior.

EJEMPLO 8 Por ejemplo, sean $v_1 = (1, 3, -1)$ y $v_2 = (2, 1, 3)$ dos vectores en $\mathbf{R}^3$. Según la discusión anterior, el menor subespacio de $\mathbf{R}^3$ que contiene a v_1 y v_2 es $\mathcal{L}(v_1, v_2)$ donde

$$\begin{aligned}\mathcal{L}(v_1, v_2) &= \{c_1(1, 3, -1) + c_2(2, 1, 3) \mid c_1, c_2 \in \mathbf{R}\}\\ &= \{(c_1 + 2c_2, 3c_1 + c_2, -c_1 + 3c_2) \mid c_1, c_2 \in \mathbf{R}\}\\ &= \{(x, y, z) \mid x = c_1 + 2c_2,\ y = 3c_1 + c_2,\ z = -c_1 + 3c_2,\ c_1, c_2 \in \mathbf{R}\}\end{aligned}$$

Se desea describir este subespacio de $\mathbf{R}^3$ en otros términos.

Obsérvese que de la última expresión para $\mathcal{L}(v_1, v_2)$ se obtiene

$$\begin{aligned} c_1 + 2c_2 &= x\\ 3c_1 + c_2 &= y\\ -c_1 + 3c_2 &= z\end{aligned}$$

que puede contemplarse como un sistema de 3 ecuaciones con 2 incógnitas c_1 y c_2, *el cual se sabe que tiene solución* (pues el vector (x, y, z) es una combinación lineal de v_1, v_2, lo que significa que existen c_1 y c_2, tales que $(x, y, z) = c_1v_1 + c_2v_2$).

Al aplicar el método de eliminación Gaussiana a este sistema se obtiene

$$\begin{bmatrix} 1 & 2 & x\\ 3 & 1 & y\\ -1 & 3 & z\end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 2 & x\\ 0 & -5 & y-3x\\ 0 & 0 & -2x+y+z\end{bmatrix}$$

Entonces, el hecho de la existencia de soluciones del sistema es equivalente a que $-2x + y + z = 0$. Es decir, el vector $(x, y, z) \in \mathbf{R}^3$ es una combinación lineal de $v_1 = (1, 3, -1)$ y $v_2 = (2, 1, 3)$ si, y sólo si $-2x + y + z = 0$. Por lo tanto, el subespacio $\mathcal{L}(v_1, v_2)$ puede ser descrito como

$$\mathcal{L}(v_1, v_2) = \{(x, y, z) \mid -2x + y + z = 0\}$$

que es un subespacio de $\mathbf{R}^3$ que representa geométricamente un plano (el plano $-2x + y + z = 0$) que pasa por el origen.

Geométricamente éste es el plano en el que se encuentran v_1 y v_2.

DEFINICIÓN 3.3 Sea V un espacio vectorial y sean $v_1, \ldots, v_n$ vectores de V. Al subespacio $\mathcal{L}(v_1, \ldots, v_n)$ se le llama *espacio generado por los vectores* $v_1, \ldots, v_n$.

En general, el subespacio $\mathcal{L}(v_1, \ldots, v_n)$ es diferente del espacio vectorial V (es decir, $\mathcal{L}(v_1, \ldots, v_n) \subsetneq V$), tal como se vio en el ejemplo $V = \mathbf{R}^3$, $v_1 = (1, 3, -1)$ y $v_2 = (2, 1, 3)$. Sin embargo, puede ocurrir que $\mathcal{L}(v_1, \ldots, v_n) = V$. En este caso se dirá que el espacio vectorial V es *generado* por los vectores $v_1, \ldots, v_n$ y se llamará a éstos los *generadores* de V. Surge, entonces, naturalmente la siguiente pregunta: ¿bajo qué condiciones (en los vectores $v_1, \ldots, v_n$) se tiene $\mathcal{L}(v_1, \ldots, v_n) = V$? La respuesta a esta pregunta se verá en la sección 5 de este capítulo.

3.2 OPERACIONES CON SUBESPACIOS

Sean W_1 y W_2 dos subespacios del espacio vectorial V. En esta subsección se considerarán los siguientes subconjuntos de V:

1) la *intersección* de W_1 y W_2

$$W_1 \cap W_2 = \{v \in V \mid v \in W_1 \text{ y } v \in W_2\}$$

2) la *unión* de W_1 y W_2

$$W_1 \cup W_2 = \{v \in V \mid v \in W_1 \text{ o } v \in W_2\}$$

3) la suma de W_1 y W_2, denotada por $W_1 + W_2$ y definida por

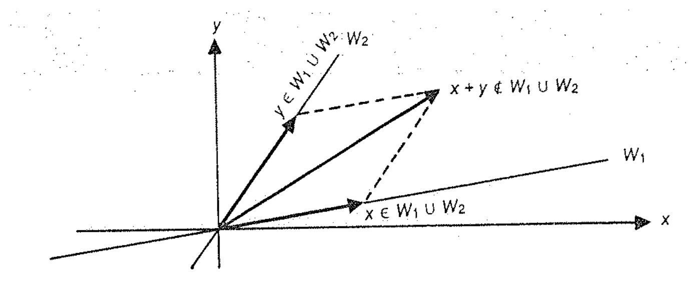
$$W_1 + W_2 = \{v \in V \mid v = w_1 + w_2,\quad w_1 \in W_1, w_2 \in W_2\}$$

Siendo W_1 y W_2 subespacios de V, es natural preguntarse si $W_1 \cap W_2$, $W_1 \cup W_2$ y $W_1 + W_2$ son también subespacios de V. A a lo largo del desarrollo de esta subsección se dará respuesta a esta pregunta.

Es inmediato verificar que $W_1 \cap W_2$ es un subespacio: es no vacío (contiene al cero) y si $x, y \in W_1 \cap W_2$ y $\lambda \in \mathbf{R}$, entonces $x, y \in W_1$ y $x, y \in W_2$. Por ser W_1 y W_2 subespacios de V, se tiene $x + y \in W_1, x + y \in W_2, \lambda x \in W_1$ y $\lambda x \in W_2$ o sea, $x + y \in W_1 \cap W_2$ y $\lambda x \in W_1 \cap W_2$, lo que prueba entonces que $W_1 \cap W_2$ es un subespacio de V.

Más aún, $W_1 \cap W_2$ puede ser contemplado también como subespacio de W_1 (o W_2). La demostración de este hecho es similar a la dada anteriormente y se deja como ejercicio para el lector.

Por otra parte la unión de dos subespacios de V no es en general un subespacio de V. La siguiente figura ilustra este hecho siendo V el espacio $\mathbf{R}_2$.



Al tomar $x \in W_1$ y $y \in W_2$, se tiene que tanto x como y están en $W_1 \cup W_2$ y, sin embargo, $x + y \notin W_1 \cup W_2$.

El siguiente teorema establece condiciones necesarias para que la unión de dos subespacios sea un subespacio (las condiciones establecidas son obviamente suficientes).

TEOREMA 3.4 Sean W_1 y W_2 dos subespacios del espacio vectorial V. Si $W_1 \cup W_2$ es subespacio de V, entonces $W_1 \subset W_2$ o $W_2 \subset W_1$.

DEMOSTRACIÓN Supóngase que $W_1 \not\subset W_2$ y pruébese entonces que $W_2 \subset W_1$. Si $W_1 \not\subset W_2$, existe un vector $y \in W_1$ tal que $y \notin W_2$. Sea $x \in W_2$. Entonces se tiene $x \in W_2 \subset W_1 \cup W_2$ y $y \in W_1 \subset W_1 \cup W_2$.

Como $W_1 \cup W_2$ es un subespacio de V, se concluye que $x + y \in W_1 \cup W_2$, es decir $x + y \in W_1$ o $x + y \in W_2$. Se afirma que $x + y \notin W_2$, pues caso contrario, ya que $x \in W_2$ se tendría que $(x + y) + (-x) \in W_2$, o sea $y \in W_2$ lo que es una contradicción a la elección de y. Entonces $x + y \in W_1$. Pero $y \in W_1$, entonces $(x + y) + (-y) \in W_1$, o sea $x \in W_1$, lo que prueba que $W_2 \subset W_1$.

Q.E.D.

La suma de los subespacios W_1 y W_2 es el conjunto de vectores $v \in V$ que se pueden escribir como la suma de un vector de W_1 y un vector de W_2. Véase que este conjunto es un subespacio de V: es no vacío (contiene al cero ¿por qué?) y si $x, y \in W_1 + W_2$ y $\lambda \in \mathbf{R}$ entonces existen $w_1, w_1' \in W_1$ y $w_2, w_2' \in W_2$ tales que $x = w_1 + w_2$ y $y = w_1' + w_2'$. Al hacer la suma de x con y se obtiene el vector $x + y = (w_1 + w_2) + (w_1' + w_2') = (w_1 + w_1') + (w_2 + w_2')$. Pero $w_1 + w_1' \in W_1$ y $w_2 + w_2' \in W_2$, pues tanto W_1 como W_2 son subespacios de V. Esto muestra que $x + y \in W_1 + W_2$. Similarmente, se tiene $\lambda x = \lambda(w_1 + w_2) = \lambda w_1 + \lambda w_2$, en donde $\lambda w_1 \in W_1$ y $\lambda w_2, \in W_2$ de modo que $\lambda x \in W_1 + W_2$ y finalmente se concluye que $W_1 + W_2$ es un subespacio de V.

El siguiente teorema dice que el subconjunto $W_1 \cup W_2$ y el subespacio $W_1 + W_2$ están íntimamente relacionados.

TEOREMA 3.5 Sean W_1 y W_2 dos subespacios del espacio vectorial V. Entonces

$$W_1 + W_2 = \mathcal{L}(W_1 \cup W_2)$$

(es decir, el espacio generado por $W_1 \cup W_2$ es $W_1 + W_2$).

DEMOSTRACIÓN Es obvio que $W_1 + W_2$ es un subespacio de V que contiene a $W_1 \cup W_2$. Entonces $\mathcal{L}(W_1 \cup W_2) \subset W_1 + W_2$ (pues $\mathcal{L}(W_1 \cup W_2)$ es el más pequeño de los subespacios de V que contiene a $W_1 \cup W_2$). Tómese ahora un vector $x = w_1 + w_2 \in W_1 + W_2$. Se tiene que $w_1 \in W_1 \subset W_1 \cup W_2$ y $w_2 \in W_2 \subset W_1 \cup W_2$. Obsérvese que en $\mathcal{L}(W_1 \cup W_2)$ se encuentran todas las combinaciones lineales (finitas) de vectores en $W_1 \cup W_2$. Entonces $x = 1 \cdot w_1 + 1 \cdot w_2 \in \mathcal{L}(W_1 \cup W_2)$, lo que muestra que $W_1 + W_2 \subset \mathcal{L}(W_1 \cup W_2)$. Por lo tanto,

$$W_1 + W_2 = \mathcal{L}(W_1 \cup W_2)$$

Q.E.D.

EJEMPLO 9 Vea un ejemplo. Sea $V = M_{2\times2}$ (espacio vectorial de las matrices de orden 2). Sea W_1 el subconjunto de $M_{2\times2}$ formado por las matrices triangulares superiores y W_2 el subconjunto formado por las matrices triangulares inferiores. Es decir,

$$W_1 = \left\{ \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \middle| \, a, b, c \in \mathbf{R} \right\} \quad \text{(matrices triangulares superiores)}$$

$$W_2 = \left\{ \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \middle| \, a, b, c \in \mathbf{R} \right\} \quad \text{(matrices triangulares inferiores)}$$

Se comprueba fácilmente que W_1 y W_2 son subespacios de $M_{2\times2}$. Sea v un vector en $W_1 \cap W_2$. Entonces v es una matriz que es a la vez triangular superior y triangular inferior. Es por tanto una matriz diagonal. Es decir,

$$W_1 \cap W_2 = \left\{ \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix} \middle| \, a, c \in \mathbf{R} \right\} \quad \text{(matrices diagonales)}$$

Sea ahora v un vector en $W_1 + W_2$. Entonces v es una matriz que se puede escribir como la suma de una matriz triangular superior y una matriz triangular inferior. Pero es claro que *toda* matriz de $M_{2\times2}$ tiene esta propiedad. Es decir,

$$W_1 + W_2 = M_{2\times2}$$

por ejemplo:

$$\underbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}_{\in M_{2\times2}} = \underbrace{\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}}_{\in W_1} + \underbrace{\begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix}}_{\in W_2}$$

Se desea llamar la atención al hecho de que los vectores v del subespacio $W_1 + W_2$ son vectores para los que existe $w_1 \in W_1$ y $w_2 \in W_2$ tales que $v = w_1 + w_2$, pero que estos vectores w_1 y w_2 *no son únicos* en general.

De hecho, regresando al ejemplo del espacio $M_{2\times2}$ se tiene

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} + \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}}_{\in W_1} + \underbrace{\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}}_{\in W_2}$$

$\vdots$
etc.

En realidad, el caso en el que ocurre que los vectores v del subespacio $W_1 + W_2$ tienen una única representación del tipo $v = w_1 + w_2$ con $w_1 \in W_1$ y $w_2 \in W_2$, es un caso muy particular que se considera ahora.

DEFINICIÓN 3.4 Sean W_1 y W_2 dos subespacios del espacio vectorial V. La suma $W_1 + W_2$ es llamada *suma directa* de W_1 y W_2, denotada por $W_1 \oplus W_2$, si cada vector en el subespacio $W_1 + W_2$ tiene una *única* representación como la suma de un vector en W_1 y un vector en W_2.

El siguiente teorema da una condición muy fácil de verificar para saber si la suma de dos subespacios de V es una suma directa.

TEOREMA 3.6 Sean W_1 y W_2 dos subespacios del espacio vectorial V. Escriba $U = W_1 + W_2$. Las dos condiciones siguientes son equivalentes

a) $U = W_1 \oplus W_2$
b) $W_1 \cap W_2 = \{0\}$

DEMOSTRACIÓN 1) $\Rightarrow$ 2). Ciertamente $W_1 \cap W_2 \subset W_1 + W_2$.
Sea $x \in W_1 \cap W_2$. Como $x \in W_1$ escríbase

$$x = x + 0 \in W_1 + W_2$$

Similarmente, como $x \in W_2$, escríbase

$$x = 0 + x \in W_1 + W_2$$

Ahora bien, el vector x como elemento de $W_1 + W_2$, tiene una única representación en este subespacio. Comparando entonces las expresiones anteriores

$$x = x + 0$$
$$x = 0 + x$$

Se concluye que $x = 0$. Es decir, $W_1 \cap W_2 = \{0\}$.

2) $\Rightarrow$ 1). Supóngase que $x \in W_1 + W_2$ tiene dos representaciones distintas, dígase

$$x = w_1 + w_2$$
$$x = w_1' + w_2'$$

Entonces

$$w_1 + w_2 = w_1' + w_2'$$

o bien,

$$w_1 - w_1' = w_2' - w_2$$

Ahora bien, $w_1 - w_1'$ pertenece a W_1 (pues w_1 y w_1' pertenecen a W_1 y éste es un subespacio de V). Similarmente, $w_2' - w_2$ pertenece a W_2. Por tanto, $w_1 - w'_1$ pertenece tanto a W_1 como a W_2 (pues es igual a $w'_2 - w_2$ que pertenece a W_2). Entonces $w_1 - w'_1 \in W_1 \cap W_2$. Pero $W_1 \cap W_2 = \{0\}$, lo que permite concluir que $w_1 = w_1'$. Un argumento similar muestra que $w_2 = w_2'$, y por tanto que la representación de $x \in W_1 + W_2$ es única.

Q.E.D.

EJEMPLO 10 Considérese por ejemplo el espacio $\mathbf{F(R)}$ (esto es, el espacio de todas las funciones reales definidas en la recta). Sean W_1 y W_2 los siguientes subconjuntos de $\mathbf{F(R)}$.

$$W_1 = \{f \in F(\mathbf{R}) \mid f(x) = f(-x)\ \forall\, x \in \mathbf{R}\}$$
$$W_2 = \{f \in F(\mathbf{R}) \mid f(x) = -f(-x)\ \forall\, x \in \mathbf{R}\}$$

A los elementos de W_1 se les llama *funciones pares* y a los de W_2 se les llama *funciones impares*.

Es un ejercicio sencillo comprobar que W_1 y W_2 son subespacios de $\mathbf{F(R)}$.

Sea f un vector cualquiera de $\mathbf{F(R)}$.

Defínanse las funciones g y h como:

$$g(x) = \tfrac{1}{2}(f(x) + f(-x)),\ h(x) = \tfrac{1}{2}(f(x) - f(-x))$$

Obsérvese que

$$g(-x) = \tfrac{1}{2}(f(-x) + f(x)) = g(x)$$
$$h(-x) = \tfrac{1}{2}(f(-x) - f(x)) = -\tfrac{1}{2}(f(x) - f(-x)) = -h(x)$$

O sea que $g \in W_1$ y $h \in W_2$.

Pero $f = g + h$, lo que muestra entonces que

$$F(\mathbf{R}) = W_1 + W_2$$

Aún más, se afirma que esta suma es directa.

En efecto, según el teorema 3.6 basta verificar que $W_1 \cap W_2 = \{0\}$ (en donde 0 denota aquí la función cero, esto es, $0: \mathbf{R} \to \mathbf{R}$, $0(x) = 0$).

Supóngase que $\varphi \in W_1 \cap W_2$. Entonces para $x \in \mathbf{R}$ se tiene

$$\underset{\text{pues } \varphi \in W_1}{\varphi(-x) \;\;=\;\; \varphi(x)} \;\;=\;\; \underset{\text{pues } \varphi \in W_2}{-\varphi(-x)}$$

de donde $\varphi(x) = 0$, esto es, φ es la función cero.

Entonces se tiene de hecho

$$F(\mathbf{R}) = W_1 \oplus W_2$$

Se ha logrado entonces presentar el espacio de todas las funciones reales definidas en la recta como la suma directa del subespacio de funciones pares y el subespacio de funciones impares.

EJERCICIOS (SECCIÓN 3, CAPÍTULO 3)

1. Demuestre que los siguientes subconjuntos de $\mathbf{R}^3$ son subespacios de este espacio vectorial.

   a) $\{(x, y, z) \mid ax + by + cz = 0\}$
   b) $\{(x, y, z) \mid x = at, y = bt, z = ct \quad t \in \mathbf{R}\}$

2. Diga si cada uno de los siguientes subconjuntos del espacio vectorial $\mathbf{R}^n$ son subespacios o no lo son

   a) $\{(x_1, x_2, \ldots, x_n) \mid x_1 + x_2 + \ldots + x_n = 0\}$
   b) $\{(x_1, x_2, \ldots, x_n) \mid x_1 = x_2 = \ldots = x_n\}$
   c) $\{(x_1, x_2, \ldots, x_n) \mid c_1x_1 + c_2x_2 + \ldots + c_nx_n = 0\}$
   d) $\{(x_1, x_2, \ldots, x_n) \mid x_1 = 0\}$
   e) $\{(x_1, x_2, \ldots, x_n) \mid x_1 + x_2 = 2\}$
   f) $\{(x_1, x_2, \ldots, x_n) \mid x_1x_2x_3 \geq 3\}$
   g) $\{(x_1, x_2, \ldots, x_n) \mid x_n = 1\}$
   h) $\{(x_1, x_2, \ldots, x_n) \mid x_1^2 + x_2^2 + \ldots + x_n^2 = 0\}$
   i) $\{(x_1, x_2, \ldots, x_n) \mid x_1 = c_1t, x_2 = c_2t, \ldots, x_n = c_nt, \quad t \in \mathbf{R}\}$

3. Explique si cada uno de los siguientes subconjuntos del espacio vectorial $\mathbf{P}_3$ son subespacios de él o no lo son

   a) $\{a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_0 + a_1 = 0\}$
   b) $\{a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_0 = a_1 = a_2 = a_3\}$
   c) $\{a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_3 = 1\}$
   d) $\{a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_0 = a_1 = a_2 = 0\}$

   e) $\{p(x) \in P_3 \mid p(1) = 0\}$
   f) $\{p(x) \in P_3 \mid p(-1) = p(1) = 0\}$
   g) $\{p(x) \in P_3 \mid p(0) = p(1) = p(2) = p(3) = 0\}$

4. Diga si los siguientes subconjuntos del espacio vectorial $M_{n \times n}$ son subespacios de él o no lo son

   a) $\{A = (a_{ij})_{i,j=1,\ldots,n} \mid a_{ij} = 0 \text{ si } i \neq j\}$ (matrices diagonales)
   b) $\{A = (a_{ij})_{i,j=1,\ldots,n} \mid a_{11} + a_{22} + \ldots + a_{nn} = 0\}$
   c) $\{A = (a_{ij})_{i,j=1,\ldots,n} \mid a_{11} + a_{12} + \ldots + a_{1n} = 0\}$
   d) $\{A = (a_{ij})_{i,j=1,\ldots,n} \mid a_{ij} = 0 \quad i, j = 1, \ldots, n\}$
   e) $\{A = (a_{ij})_{i,j=1,\ldots,n} \mid a_{ij} = 0 \quad \text{si } i > j\}$ (matrices triangulares superiores)
   f) $\{A = (a_{ij})_{i,j=1,\ldots,n} \mid a_{ij} = 0 \quad \text{si } i < j\}$ (matrices triangulares inferiores)
   g) $\{A \in M_{n \times n} \mid AB = BA, B \text{ es una matriz fija de } M_{n \times n}\}$
   h) $\{A \in M_{n \times n} \mid A = A^t\}$ (matrices simétricas)
   i) $\{A \in M_{n \times n} \mid A = -A^t\}$ (matrices antisimétricas)
   j) $\{A \in M_{n \times n} \mid A \text{ es inversible}\}$

5. Considere el espacio vectorial $F(I)$ de las funciones reales definidas en el subconjunto I de $\mathbf{R}$. Se dice que la función $f \in F(I)$ es de clase C^k (k un número natural) si f es una función derivable k veces y la función $f^{(k)}$ es una función continua. Demuestre que el subconjunto de $F(I)$ formado por las funciones de clase C^k (denotado por $C^k(I)$ es un subespacio de $F(I)$.

② 6. Considere el espacio vectorial $C^2(\mathbf{R})$ —véase ejercicio anterior. Sean a, b y c tres constantes fijas. Demuestre que el conjunto de funciones $f \in C^2(\mathbf{R})$ tales que $af''(x) + bf'(x) + cf(x) = 0$, es un subespacio de $C^2(\mathbf{R})$. Más generalmente, en el espacio vectorial $C^n(\mathbf{R})$, el conjunto de funciones f tales que

$$a_nf^{(n)}(x) + a_{n-1}f^{(n-1)}(x) + \ldots + a_1f'(x) + a_0f(x) = 0$$

en donde $a_0, a_1, \ldots, a_n$ son constantes fijas, es un subespacio de $C^n(\mathbf{R})$.

7. Describa explícitamente los espacios solución de los siguientes sistemas homogéneos de ecuaciones lineales:

   a) $\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ 2x_1 + 2x_2 + 2x_3 &= 0 \\ 3x_1 - x_2 + x_3 &= 0 \end{aligned}$

   b) $\begin{aligned} 3x_1 - x_2 - 2x_3 &= 0 \\ x_1 + x_2 - 2x_3 &= 0 \\ x_1 - 3x_2 + 2x_3 &= 0 \end{aligned}$

   c) $\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 2x_1 + 2x_2 - 2x_3 &= 0 \end{aligned}$

   d) $\begin{aligned} x_1 - x_2 + x_3 - x_4 &= 0 \\ 2x_1 + x_2 - x_3 + 2x_4 &= 0 \\ x_1 + 3x_2 - 2x_3 + x_4 &= 0 \end{aligned}$

8. Sea A una matriz inversible de orden n. Describa el espacio solución del sistema homogéneo de ecuaciones lineales $AX = 0$.

9. Sea A una matriz cuadrada de orden n. Considere el sistema no homogéneo de ecuaciones lineales $AX = B$. ¿Es el conjunto solución de este sistema un subespacio de $\mathbf{R}^n$?

10. Demuestre que el vector $(2, 5)$ de $\mathbf{R}^2$ es una combinación lineal de los vectores $v_1 = (1, 3)$ y $v_2 = (2, -1)$.

11. Demuestre que el vector $\begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix}$ de $M_{2\times 2}$ es una combinación lineal de los vectores

$$v_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad v_2 = \begin{bmatrix} -1 & 3 \\ 0 & 4 \end{bmatrix} \quad v_3 = \begin{bmatrix} 0 & 1 \\ 3 & -1 \end{bmatrix} \quad v_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

12. Determine cuáles de los siguientes vectores de P_2 son combinaciones lineales de los vectores $p_1(x) = x + x^2$ y $p_2(x) = 1 + x$
   a) $p(x) = -2 - x + x^2$
   b) $p(x) = 3 + 2x - x^2$
   c) $p(x) = 1 - 4x + x^2$
   d) $p(x) = 10 - 2x + 8x^2$
13. Determine el menor de los subespacios de $\mathbf{R}^2$ en el que se encuentran los vectores $v_1 = (1, 3)$ y $v_2 = (-1, 2)$.
14. Determine el menor de los subespacios de $M_{2\times 2}$ en el que se encuentran los vectores

$$v_1 = \begin{bmatrix} 1 & 3 \\ 8 & 1 \end{bmatrix} \quad v_2 = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}$$

15. Demuestre que el menor subespacio de $\mathbf{R}^3$ en el que se encuentran los vectores $v_1 = (1, 0, -1)$, $v_2 = (0, 2, -2)$ y $v_3 = (0, -2, 2)$ es el plano $x + y + z = 0$.
16. Demuestre que el menor subespacio de $\mathbf{R}^3$ en el que se encuentran los vectores $v_1 = (4, -2, 6)$ y $v_2 = (-2, 1, -3)$ es la recta $x = 2t$, $y = -t$, $z = 3t$ $(t \in \mathbf{R})$.
17. Demuestre que no existe subespacio propio de $\mathbf{R}^3$ en el que se encuentren los vectores $v_1 = (1, 1, 1)$, $v_2 = (1, 1, 0)$ y $v_3 = (1, 0, 0)$.
18. ¿Cuál es el menor subespacio de $\mathbf{R}^n$ en el que se encuentra el vector cero de este espacio?
19. Describa explícitamente los siguientes subespacios de $\mathbf{R}^3$:
   a) $\mathcal{L}((1, 2, -1))$
   b) $\mathcal{L}((1, 1, 0), (-3, 3, 0))$
   c) $\mathcal{L}((2, 1, 4), (3, -1, 0))$
   d) $\mathcal{L}((5, 1, 2), (6, 1, 3), (-1, 0, -1))$
   e) $\mathcal{L}((2, 1, 2), (-1, 0, 4), (3, 1, 4))$
   f) $\mathcal{L}((1, 1, 2), (4, 2, 1), (3, 1, -1), (-1, 1, 5))$
   g) $\mathcal{L}((1, 0, 0), (0, 1, 0), (0, 0, 1))$
   h) $\mathcal{L}((1, 1, 1), (2, 2, 2), (-3, -3, -3), (4, 4, 4), (-5, -5, -5))$
20. Considere el subespacio S de $M_{2\times 2}$ generado por los vectores

$$v_1 = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \quad v_2 = \begin{bmatrix} 2 & -1 \\ 0 & 4 \end{bmatrix}$$

   Determine cuáles de los siguientes vectores pertenecen a S:

   a) $\begin{bmatrix} 2 & -15 \\ 0 & 8 \end{bmatrix}$ c) $\begin{bmatrix} 3 & 1 \\ 0 & -1 \end{bmatrix}$ e) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

   b) $\begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix}$ d) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ f) $\begin{bmatrix} 1 & -4 \\ 0 & 3 \end{bmatrix}$

21. Considere el subespacio S de $F(\mathbf{R})$ generado por los vectores $\operatorname{sen}^2 x$ y $\cos^2 x$. Demuestre que los vectores 1 y $\cos 2x$ de $F(\mathbf{R})$ pertenecen a S.

22. Demuestre que el espacio vectorial $\mathbf{R}^3$ es generado por cada uno de los siguientes conjuntos de vectores:
   a) $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$, $v_3 = (0, 0, 1)$
   b) $v_1 = (1, 1, 1)$, $v_2 = (2, 3, -5)$, $v_3 = (0, 1, 0)$
   c) $v_1 = (2, 3, 1)$, $v_2 = (3, 4, 2)$, $v_3 = (0, 1, 3)$, $v_4 = (1, 3, -1)$
23. Pruebe que el espacio vectorial $\mathbf{R}^4$ no es generado por los siguientes conjuntos de vectores:
   a) $v_1 = (2, 1, 3, 1)$, $v_2 = (3, 1, 1, 1)$
   b) $v_1 = (0, 1, 3, 2)$, $v_2 = (1, 3, -1, 4)$, $v_3 = (0, 1, 1, 0)$
   c) $v_1 = (3, 1, 1, 1)$, $v_2 = (1, 2, -1, 1)$, $v_3 = (5, 5, -1, 3)$, $v_4 = (3, 1, 0, 0)$
24. Demuestre que el espacio vectorial P_n es generado por el conjunto de $n + 1$ vectores $1, x, x^2, \ldots, x^n$.
25. Pruebe que el espacio vectorial P_n no es generado por los vectores $1, (x + 1), (x + 1)^2, \ldots, (x + 1)^{n-1}$.
26. Demuestre que el espacio vectorial $\mathbf{R}^n$ no es generado por los vectores $(1, 1, \ldots, 1, 1, 0), (1, 1, \ldots, 1, 0, 0), \ldots, (1, 0, 0, \ldots, 0, 0)$.
27. Considere el subespacio de $\mathbf{R}^2$ dado por

$$S = \{(x, y, z) \mid 2x - 3y + 5z = 0\}$$

   Encuentre dos vectores v_1 y v_2 en S que generen este subespacio. ¿Puede S ser generado por un solo vector?
28. Cierto o falso
   a) El espacio que genera un vector en $\mathbf{R}^3$ es una recta que pasa por el origen. ________
   b) Un plano que pasa por el origen en $\mathbf{R}^3$ es un subespacio que es generado por 2 vectores cualesquiera sobre el plano. ________
   c) El espacio que generan dos vectores no nulos en $\mathbf{R}^3$ es un plano que pasa por el origen. ________
   d) Un solo vector en $\mathbf{R}^3$ no puede generar un plano. ________
   e) El espacio que generan 3 o más vectores en $\mathbf{R}^3$ es todo $\mathbf{R}^3$. ________
   f) $\mathbf{R}^3$ no puede ser generado por 2 vectores. ________
29. Sea V un espacio vectorial y sea S un subconjunto de V. Considere el conjunto de todos los subespacios de V que contienen a S

$$T = \{W \subseteq V \mid W \text{ es subespacio de } V \text{ y } W \supseteq S\}$$

   Observe que T no es vacío pues V mismo está en él.
   a) Demuestre que la intersección de *todos* los elementos de T es un subespacio de V que contiene a S. Es decir, demuestre que

$$\tilde{W} = \bigcap_{W \in T} W \in T$$

   b) Demuestre que $\tilde{W}$ es el menor de los subespacios (en el sentido de contención de conjuntos) de V que contienen a S. Concluya, entonces, que este subespacio $\tilde{W}$ es el subespacio de todas las combinaciones lineales finitas de vectores de S. Es decir, concluya que $W = \mathcal{L}(S)$.
   c) Describa el subespacio generado por el conjunto vacío.

30. Sean S_1 y S_2 dos subespacios del espacio vectorial $\mathbf{R}^3$. Al usar argumentos geométricos, describa todas las posibilidades para el subespacio $S_1 \cap S_2$.
31. Sean S_1 y S_2 dos subespacios del espacio vectorial $\mathbf{R}^3$. Al usar argumentos geométricos, describa todas las posibilidades para el subespacio $S_1 + S_2$.
32. Sean S_1 y S_2 los siguientes subespacios de $\mathbf{R}_3$

$$S_1 = \{(x, y, z) \mid x + y + z = 0\}$$
$$S_2 = \{(x, y, z) \mid 2x - y + z = 0\}$$

a) Describa explícitamente el subespacio $S_1 \cap S_2$.
b) Con un ejemplo numérico, muestre que $S_1 \cup S_2$ no es un subespacio de $\mathbf{R}^3$.
c) Demuestre que el subespacio $S_1 + S_2$ es todo $\mathbf{R}^3$.

33. Interprete geométricamente el teorema 3.4 en el caso en el que V es el espacio vectorial $\mathbf{R}^3$.
34. Considere los subespacios W_1 y W_2 de $\mathbf{R}^2$ dados por

$$W_1 = \mathcal{L}((1, 0)) \qquad W_2 = \mathcal{L}((0, 1))$$

Demuestre que $\mathcal{L}(W_1 \cup W_2) = \mathbf{R}^2$. Interprete geométricamente este hecho.
35. Generalice la definición de unión, suma y suma directa de dos subespacios dada en esta sección al caso de unión, suma y suma directa de n subespacios de un espacio vectorial.
36. Generalice el teorema 3.5: Si $W_1, W_2, \ldots, W_n$ son n subespacios del espacio vectorial V, entonces

$$W_1 + W_2 + \ldots + W_n = \mathcal{L}(W_1 \cup W_2 \cup \ldots \cup W_n)$$

37. Considere los subespacios W_1 y W_2 del ejercicio 34. Demuestre que $\mathbf{R}^2$ es la suma directa de W_1 y W_2.
38. Considere los subespacios $W_1, W_2, \ldots, W_n$ de $\mathbf{R}^n$ dados por

$$W_i = \mathcal{L}((0, 0, \ldots, \underbrace{1}_{i}, \ldots, 0)) \qquad i = 1, 2, \ldots, n$$

Demuestre que

$$\mathbf{R}^n = W_1 \oplus W_2 \oplus \ldots \oplus W_n$$

4. DEPENDENCIA E INDEPENDENCIA LINEAL

DEFINICIÓN 4.1 Sea V un espacio vectorial y sean $v_1, v_2, \ldots, v_n$ vectores de V. Se dice que estos vectores son *linealmente independientes* si se cumple

$$c_1v_1 + c_2v_2 + \ldots + c_nv_n = 0 \Rightarrow c_1 = c_2 = \ldots = c_n = 0$$

Caso contrario, se dirá que tales vectores son *linealmente dependientes*, es decir, si existen escalares $c_1, c_2, \ldots, c_n$, al menos uno de ellos distinto de cero, tales que $c_1v_1 + c_2v_2 + \ldots + c_nv_n = 0$.*

*No se hará distinción alguna entre la afirmación "los vectores $v_1, v_2, \ldots, v_n$ son linealmente independientes (dependientes)" y "el conjunto $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente (dependiente)".

EJEMPLO 1 Por ejemplo, se afirma que los vectores $v_1 = (1, 3, 2)$, $v_2 = (-1, 2, 5)$ y $v_3 = (0, 1, 4)$ en $\mathbf{R}^3$, son linealmente independientes.

Para ver la validez de la afirmación, se tiene que considerar la combinación lineal de ellos (igualada al vector cero de $\mathbf{R}^3$)

$$c_1(1, 3, 2) + c_2(-1, 2, 5) + c_3(0, 1, 4) = (0, 0, 0)$$

y mostrar que esto implica que los 3 escalares c_1, c_2 y c_3 son iguales a cero.

La combinación lineal anterior puede reescribirse como (realizando las operaciones indicadas en el lado izquierdo de la igualdad)

$$(c_1 - c_2, 3c_1 + 2c_2 + c_3, 2c_1 + 5c_2 + 4c_3) = (0, 0, 0)$$

de donde, igualando las coordenadas correspondientes de estos dos vectores de $\mathbf{R}^3$ se obtiene

$$c_1 - c_2 = 0$$
$$3c_1 + 2c_2 + c_3 = 0$$
$$2c_1 + 5c_2 + 4c_3 = 0$$

el cual es un sistema homogéneo de 3 ecuaciones lineales para las incógnitas c_1, c_2, c_3. El determinante de la matriz de este sistema es:

$$det\begin{bmatrix} 1 & -1 & 0 \\ 3 & 2 & 1 \\ 2 & 5 & 4 \end{bmatrix} = 13 \neq 0$$

lo que dice, entonces, que la matriz del sistema es inversible y que por lo tanto el sistema homogéneo correspondiente tiene solamente la solución trivial $c_1 = c_2 = c_3 = 0$. Esto prueba la afirmación inicial.

EJEMPLO 2 Por otra parte, los vectores v_1, v_2, v_3 en $M_{2\times2}$ dados por

$$v_1 = \begin{bmatrix} -1 & 16 \\ -4 & 13 \end{bmatrix} \quad v_2 = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} \quad v_3 = \begin{bmatrix} -1 & 2 \\ -2 & 1 \end{bmatrix}$$

son linealmente independientes.

En efecto, vea que existen escalares c_1, c_2 y c_3 no nulos (al menos uno de ellos no nulo) tales que

$$c_1\begin{bmatrix} -1 & 16 \\ -4 & 13 \end{bmatrix} + c_2\begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} + c_3\begin{bmatrix} -1 & 2 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Al realizar las operaciones indicadas se obtiene

$$\begin{bmatrix} -c_1 + 2c_2 - c_3 & 16c_1 + 3c_2 + 2c_3 \\ -4c_1 + 3c_2 - 2c_3 & 13c_1 + 4c_2 + c_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

o sea, se tiene el sistema

$$-c_1 + 2c_2 - c_3 = 0$$
$$16c_1 + 3c_2 + 2c_3 = 0$$
$$-4c_1 + 3c_2 - 2c_3 = 0$$
$$13c_1 + 4c_2 + c_3 = 0$$

Al proceder por el método de eliminación Gaussiana se obtiene

$$\begin{bmatrix} -1 & 2 & -1 \\ 16 & 3 & 2 \\ -4 & 3 & -2 \\ 13 & 4 & 1 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 1/5 \\ 0 & 1 & -2/5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

de modo que el sistema tiene por soluciones $c_1 = -1/5\, t$, $c_2 = 2/5\, t$, $c_3 = t$, $t \in \mathbb{R}$.

Al poner $t = 5$ se obtiene en particular $c_1 = -1$, $c_2 = 2$, $c_3 = 5$ tales que

$$(-1)\begin{bmatrix} -1 & 16 \\ -4 & 13 \end{bmatrix} + 2\begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} + 5\begin{bmatrix} -1 & 2 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad (4.1)$$

lo que muestra, entonces, que v_1, v_2, v_3, son vectores linealmente dependientes.

Observe en el ejemplo anterior que el vector v_1 puede escribirse como una combinación lineal de los vectores v_2 y v_3, es decir, se tiene

$$\begin{bmatrix} -1 & 16 \\ -4 & 13 \end{bmatrix} = 2\begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} + 5\begin{bmatrix} -1 & 2 \\ -2 & 1 \end{bmatrix} \qquad (4.2)$$

Recíprocamente, suponga que se parte de la relación (4.2), la cual muestra al vector v_1 como una combinación lineal de v_2 y v_3. Escribiendo esta relación en la forma (4.1), se ve entonces que los vectores v_1, v_2, v_3 son linealmente dependientes.

Éste es un hecho general para vectores linealmente dependientes que se enuncia y se demuestra a continuación.

TEOREMA 4.1

> Los n vectores $v_1, v_2, \ldots, v_n$ $(n \geq 2)$ del espacio vectorial V son linealmente dependientes si, y sólo si al menos uno de ellos puede escribirse como combinación lineal de los $n - 1$ vectores restantes.

DEMOSTRACIÓN Si los n vectores $v_1, v_2, \ldots, v_n$ son linealmente dependientes, según la definición de dependencia lineal existen escalares $c_1, c_2, \ldots, c_n$ no todos nulos tales que

$$c_1v_1 + c_2v_2 + \ldots + c_nv_n = 0 \qquad (4.3)$$

Sin pérdida de generalidad se puede suponer que $c_1 \neq 0$ (si esto no fuese cierto, se reenumeran los vectores). Entonces la expresión (4.3) puede reescribirse como:

$$v_1 = \left(-\frac{c_2}{c_1}\right)v_2 + \ldots + \left(-\frac{c_n}{c_1}\right)v_n$$

lo que muestra que el vector v_1 es una combinación lineal de los $n - 1$ vectores $v_2, v_3, \ldots, v_n$.

Recíprocamente, supóngase que al menos uno de los n vectores $v_1, v_2, \ldots, v_n$ puede escribirse como combinación lineal de los $n - 1$ vectores restantes.

Dígase que v_1 es tal vector. Se tiene entonces que existen escalares $d_2, \ldots, d_n$ tales que

$$v_1 = d_2v_2 + \ldots + d_nv_n$$

Reescribiendo esta expresión como

$$v_1 - d_2v_2 - \ldots - d_nv_n = 0$$

se ve que se satisface (4.3) con $c_1 = 1 \neq 0$ y $c_i = -d_i$, $i = 2, 3, \ldots, n$. O sea, que los vectores $v_1, v_2, \ldots, v_n$ son linealmente dependientes.

Q.E.D.

El siguiente teorema recoge algunas consecuencias inmediatas de la definición de independencia y dependencia lineal de vectores. Su demostración se deja como ejercicio para el lector.

TEOREMA 4.2

> Sean $v_1, v_2, \ldots, v_n$, n vectores del espacio vectorial V.
>
> 1) Si $v_i = v_j$ para $i \neq j$, $i,j = 1, 2, \ldots, n$, entonces estos vectores son linealmente dependientes.
> 2) Si existe i, $1 \leq i \leq n$, tal que $v_i = 0$, entonces estos vectores son linealmente dependientes.
> 3) Si el conjunto $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente, entonces cualquier subconjunto de él es linealmente independiente.
> 4) Si el conjunto $\{v_1, v_2, \ldots, v_n\}$ es linealmente dependiente, entonces para cualesquiera vectores $w_1, w_2, \ldots, w_k$ en V, el conjunto $\{v_1, v_2, \ldots, v_n, w_1, \ldots, w_k\}$ es linealmente dependiente.

Véanse algunos ejemplos más.

Considérese el espacio vectorial $\mathbf{R}^n$. Sean $v_1, v_2, \ldots, v_m$, m vectores de este espacio. Se afirma que si $m > n$, los vectores $v_1, v_2, \ldots, v_m$, cualesquiera que éstos sean, son linealmente dependientes.

En efecto, escríbase

$$v_i = (a_{1i}, a_{2i}, \ldots, a_{ni}) \qquad i = 1, 2, \ldots, m$$

Considérese entonces la combinación lineal

$$c_1v_1 + c_2v_2 + \ldots + c_mv_m = 0 \tag{4.4}$$

Al realizar las operaciones indicadas en el lado izquierdo de esta expresión e igualando las correspondientes coordenadas de los vectores de la izquierda y derecha en ella, se obtiene

$$a_{11}c_1 + a_{12}c_2 + \ldots + a_{1m}c_m = 0$$
$$a_{21}c_1 + a_{22}c_2 + \ldots + a_{2m}c_m = 0$$
$$\cdots$$
$$a_{n1}c_1 + a_{n2}c_2 + \ldots + a_{nm}c_m = 0$$

Éste es un sistema homogéneo de n ecuaciones lineales con m incógnitas $c_1, c_2, \ldots, c_m$.

Como $m > n$, el sistema tiene soluciones no triviales. Es decir, existen valores de $c_1, c_2, \ldots, c_m$ *no todos nulos* tales que satisfacen (4.4). Esto significa precisamente que los vectores $v_1, v_2, \ldots, v_m$ son linealmente dependientes.

Obsérvese que de este ejemplo no se puede concluir nada acerca de la independencia o dependencia lineal de m vectores en $\mathbf{R}^n$ cuando $m \leq n$. De hecho, se ve en el ejemplo al inicio de esta sección el caso en el que 3 vectores en $\mathbf{R}^3$ son linealmente independientes, sin embargo, se puede tener también 3 vectores en $\mathbf{R}^3$ linealmente dependientes, por ejemplo los vectores $v_1 = (1, 3, 1)$, $v_2 = (0, -5, 2)$, $v_3 = (2, 1, 4)$, cuya dependencia lineal se sigue del hecho de que $v_2 = v_3 - 2v_1$. Lo que el ejemplo anterior nos asegura es que en $\mathbf{R}^3$ no se puede tener más de 3 vectores linealmente independientes.

En el caso particular de n vectores en $\mathbf{R}^n$, se tiene una manera sencilla de detectar la independencia o dependencia lineal de éstos. El resultado se establece en el siguiente teorema.

TEOREMA 4.3 Sean $v_j = (a_{1j}, a_{2j}, \ldots, a_{nj})$ $j = 1, 2, \ldots, n$, n vectores en $\mathbf{R}^n$. Considérese la matriz A de orden n $A = (a_{ij})_{i,j=1,2,\ldots,n}$ (esto es, A es la matriz cuya j-ésima columna está constituida por las coordenadas del j-ésimo vector v_j, $j = 1, 2, \ldots, n$). Entonces los vectores $v_1, v_2, \ldots, v_n$ son linealmente dependientes si, y sólo si $\det A = 0$ (o equivalentemente, si, y sólo si A es una matriz no inversible).

DEMOSTRACIÓN Supóngase que los vectores $v_1, v_2, \ldots, v_n \in \mathbf{R}^n$ son linealmente dependientes. Por el teorema (4.1) se puede escribir alguno de estos vectores como combinación lineal de los restantes.

Sin pérdida de generalidad, dígase que

$$v_1 = d_2v_2 + \ldots + d_nv_n \tag{4.5}$$

para ciertos escalares $d_2, \ldots d_n$. En la matriz A, realícense las siguientes operaciones en sus columnas: sustitúyase la primera columna por ella misma menos d_j veces la j-ésima columna ($j = 2, 3, \ldots, n$). Entonces en vista de (4.5) la matriz A' así obtenida tendrá su primera columna constituida toda de ceros.

Pero entonces

$$\det A = \det A' = 0$$

En forma recíproca, supóngase que $\det A = 0$. Considérese el sistema homogéneo de ecuaciones $AX = 0$. Este sistema tiene entonces soluciones no triviales. Es decir, existe

$$X_0 = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

tal que $AX_0 = 0$ y no todos los escalares $c_1, c_2, \ldots, c_n$ son iguales a cero. Al escribir explícitamente la expresión $AX_0 = 0$ se obtiene

$$\sum_{j=1}^{n} a_{ij}c_j = 0 \qquad i = 1, 2, \ldots, n$$

la cual es equivalente a

$$c_1v_1 + c_2v_2 + \ldots + c_nv_n = 0$$

Como no todos los escalares $c_1, c_2, \ldots, c_n$ son cero, esto significa que los vectores $v_1, v_2, \ldots, v_n$ son linealmente dependientes.

Q.E.D.

EJEMPLO 3 Por ejemplo, los vectores $v_1 = (2, 3, 5, -1)$, $v_2 = (0, 3, -1, 2)$, $v_3 = (0, 0, 4, 8)$ y $v_4 = (0, 0, 0, 5)$ en $\mathbf{R}^4$ son linealmente independientes puesto que

$$\det \begin{bmatrix} 2 & 0 & 0 & 0 \\ 3 & 3 & 0 & 0 \\ 5 & -1 & 4 & 0 \\ -1 & 2 & 8 & 5 \end{bmatrix} = (2)(3)(4)(5) = 120 \neq 0$$

Se desearía ahora resaltar el aspecto geométrico de la dependencia lineal de vectores en el espacio $\mathbf{R}^3$ (en donde se puede "ver" a los vectores). El hecho de que 3 vectores* en $\mathbf{R}^3$ sean linealmente dependientes se refleja geométricamente en la posición relativa entre estos vectores. Más precisamente, se tiene el siguiente resultado que nos da la versión geométrica de la dependencia lineal de 3 vectores en $\mathbf{R}^3$.

TEOREMA 4.4

> Sean $v_j = (a_{1j}, a_{2j}, a_{3j}), j = 1, 2, 3$ vectores en $\mathbf{R}^3$. Los vectores v_1, v_2, v_3 son linealmente dependientes si, y sólo si son coplanares, es decir, si los 3 se encuentran en un mismo plano (que pasa por el origen).

DEMOSTRACIÓN Recuérdese que la ecuación de un plano que pasa por el origen es:

$$Ax + By + Cz = 0$$

en donde A, B y C no son simultáneamente cero.

El hecho de que el vector $v_j = (a_{1j}, a_{2j}, a_{3j})$ $1 \leq j \leq 3$ se encuentra *en el plano* $Ax + By + Cz = 0$ significa que

$$Aa_{1j} + Ba_{2j} + Ca_{3j} = 0 \qquad 1 \leq j \leq 3$$

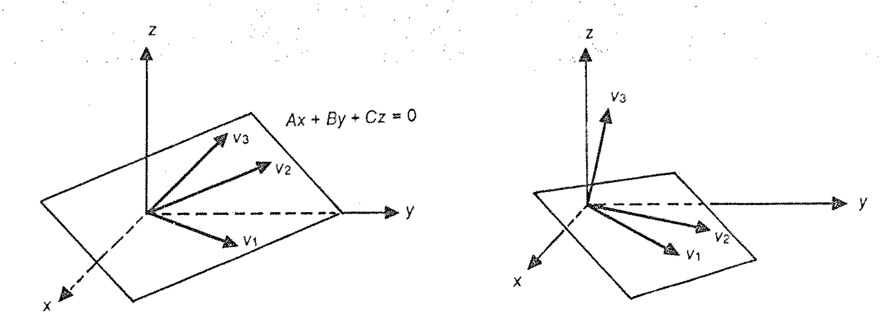
Supóngase entonces que los 3 vectores v_1, v_2, v_3 se encuentran en el plano $Ax + By + Cz = 0$. Esto significa que se tienen las ecuaciones

$$\begin{aligned} Aa_{11} + Ba_{21} + Ca_{31} &= 0 \\ Aa_{12} + Ba_{22} + Ca_{32} &= 0 \\ Aa_{13} + Ba_{23} + Ca_{33} &= 0 \end{aligned} \qquad (4.6)$$

Véanse estas ecuaciones como un sistema homogéneo de 3 ecuaciones lineales con coeficientes a_{ij}, $i, j = 1, 2, 3$. Obsérvese que el determinante de la matriz del sistema es *igual a cero* (pues se sabe que al menos uno de los valores A, B o C es distinto de cero, esto es, que el sistema tiene soluciones no triviales). Entonces por el teorema (4.3) los vectores v_1, v_2, v_3 son linealmente dependientes. En forma recíproca, supóngase que los vectores v_1, v_2, v_3 son linealmente dependientes. Por el teorema (4.3) el determinante del sistema (4.6) es igual a cero, y por tanto, existen A, B, C, al menos una de ellas distinta de cero que satisfacen cada una de las 3 ecuaciones del sistema. Esto significa que los 3 vectores v_1, v_2, v_3 se encuentran en el plano $Ax + By + Cz = 0$.

Q.E.D.

*Al usar el teorema (4.1) y la discusión geométrica de la subsección 2.1, es fácil darse cuenta del aspecto geométrico de la dependencia lineal de 2 vectores en $\mathbf{R}^2$ y $\mathbf{R}^3$. Esto se deja como ejercicio para el lector.



Véase ahora un par de ejemplos sobre independencia lineal en otros espacios vectoriales distintos de $\mathbf{R}^n$.

EJEMPLO 4

Considere en el espacio vectorial P_n los $n + 1$ vectores $v_1 = 1, v_2 = x, \ldots, v_{n+1} = x^n$. Se afirma que estos vectores son linealmente independientes.

En efecto, al escribir

$$c_1 + c_2x + \ldots + c_{n+1}x^{n+1} = 0 \qquad (4.7)$$

se puede ver fácilmente que $c_1 = c_3 = \ldots = c_n = 0$: observe que el símbolo 0 del lado izquierdo de esta última expresión se refiere al cero del espacio vectorial P_n, esto es, al polinomio 0. Este polinomio es (como ya se había visto en 2.2) aquel que tiene todos sus coeficientes iguales a 0. Es decir, se puede reescribir (4.7) como

$$c_1 + c_2x + \ldots + c_{n+1}x^{n+1} = 0 + 0 \cdot x + \ldots + 0 \cdot x^{n+1}$$

Entonces, por la definición de igualdad entre elementos del espacio P_n (véase nota al pie de la página anterior), se tiene $c_1 = c_2 = \ldots = c_{n+1} = 0$.

EJEMPLO 5

Como último ejemplo, sea V el espacio vectorial $C(\mathbf{R})$ de todas las funciones reales continuas definidas en la recta. Investigue la dependencia lineal de los vectores $v_1 = e^x$, $v_2 = \operatorname{sen} x$ y $v_3 = \cos x$ en $C(\mathbf{R})$.

Escríbase

$$c_1e^x + c_2 \operatorname{sen} x + c_3 \cos x = 0 \qquad (4.8)$$

Nuevamente debe observarse que el símbolo 0 del lado izquierdo de esta expresión se refiere a la función continua cero, esto es, la función $0: \mathbf{R} \to \mathbf{R}$ tal que $0(x) = 0\ \forall\, x \in \mathbf{R}$.

Al aprovechar entonces el hecho de que el lado izquierdo de (4.8) debe ser igual a cero (al número real 0) para cualquier $x \in \mathbf{R}$, dé 3 valores distintos a x para así generar un sistema de 3 ecuaciones (una por cada valor asignado a x) con las incógnitas c_1, c_2 y c_3.

Por ejemplo, al poner $x = 0, \pi/2, \pi$ se obtienen las ecuaciones

$$c_1 + c_3 = 0$$
$$e^{\pi/2}c_1 + c_2 = 0$$
$$e^{\pi}c_1 - c_3 = 0$$

El determinante de la matriz de este sistema vale $-1-e^{\pi} \neq 0$. Por tanto, el sistema tiene sólo la solución trivial $c_1 = c_2 = c_2 = 0$ y entonces se concluye que los vectores e^x, sen x y cos x son linealmente independientes.

NOTA: La definición de dependencia e independencia lineal dada al inicio de esta sección para un subconjunto finito de vectores de un espacio vectorial V, puede ser ampliada al caso de un subconjunto S de V no necesariamente finito. En este caso, se dice que el subconjunto S de V es linealmente independiente si cada subconjunto finito de S lo es. Caso contrario se dice que S es linealmente dependiente.

EJERCICIOS (SECCIÓN 4, CAPÍTULO 3)

1. Demuestre el teorema 4.2.
2. Pruebe que los vectores $v_1 = (a, b)$ y $v_2 = (c, d)$ en $\mathbf{R}^2$ son linealmente independientes si, y sólo si $ad - bc \neq 0$.
3. Determine cuáles de los siguientes conjuntos de vectores en $\mathbf{R}^3$ son linealmente independientes

   a) $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$, $v_3 = (0, 0, 1)$
   b) $v_1 = (2, 3, 5)$, $v_2 = (3, 1, 4)$
   c) $v_1 = (1, 3, 8)$, $v_2 = (1, 2, -5)$, $v_3 = (3, 7, -2)$
   d) $v_1 = (1, 2, 1)$
   e) $v_1 = (3, 1, 4)$, $v_2 = (2, 1, 3)$, $v_3 = (0, 0, 0)$
   f) $v_1 = (1, 1, 2)$, $v_2 = (3, 1, 1)$, $v_3 = (1, 1, 2)$
   g) $v_1 = (1, 3, 1)$, $v_2 = (2, 6, 2)$
   h) $v_1 = (1, 2, 1)$, $v_2 = (3, 1, 1)$, $v_3 = (0, 3, 4)$, $v_4 = (1, 3, 2)$
   i) $v_1 = (1, 3, 1)$, $v_2 = (2, 1, -1)$, $v_3 = (0, 3, 2)$
   j) $v_1 = (2, 1, 3)$, $v_2 = (1, 2, -1)$, $v_3 = (4, 5, 1)$
4. ¿Cierto o falso? En un espacio vectorial V cualquier conjunto constituido por un solo vector es linealmente independiente.
5. ¿Bajo qué condiciones los n vectores de $\mathbf{R}^n$

$$v_1 = (a_{11}, a_{12}, \ldots, a_{1n})$$
$$v_2 = (0, a_{22}, \ldots, a_{2n})$$
$$\cdots$$
$$v_n = (0, 0, \ldots, a_{nn})$$

son linealmente independientes?

6. Determine el valor de a de tal modo que los vectores $v_1 = (1, 2, 0, 0)$, $v_2 = (-1, 3, a, 1)$, $v_3 = (4, 1, 1, 1)$ y $v_4 = (3, 0, 0, 2)$ sean linealmente dependientes. Con tal valor de a, verifique que se satisface el teorema 4.1.
7. Demuestre que, independientemente del valor de a, los vectores de $\mathbf{R}^4$ $v_1 = (1, 3, 0, -1)$, $v_2 = (1, a, 3, 2)$, $v_3 = (1, 4, 0, -1)$ y $v_4 = (5, 1, 1, 8)$ son linealmente independientes.
8. ¿Cuál es el aspecto geométrico de dos vectores en $\mathbf{R}^2$ linealmente independientes, en $\mathbf{R}^3$?
9. Compruebe que los vectores de $\mathbf{R}^3$ $v_1 = (3, 1, 1)$, $v_2 = (-1, 2, -3)$ y $v_3 = (8, 5, 0)$ son linealmente dependientes. Encuentre la ecuación del plano en el que estos vectores se encuentran.
10. Al usar solamente argumentos geométricos (teorema 4.4), demuestre que los vectores de $\mathbf{R}^3$ $v_1 = (3, -25, 0)$, $v_2 = (87, 49, 0)$ y $v_3 = (370, 520, 1)$ son linealmente independientes.
11. Compruebe la siguiente generalización del teorema 4.4: Sean

$$v_j = (a_{1j}, a_{2j}, a_{3j}, a_{4j}) \qquad j = 1, 2, 3, 4$$

cuatro vectores en $\mathbf{R}^4$. Estos vectores son linealmente dependientes si, y sólo si todos ellos se encuentran en el hiperplano que pasa por el origen

$$Ax + By + Cx + Du = 0$$

12. Demuestre que los cuatro vectores de P_3: $1, x + 2, (x + 2)^2, (x + 2)^3$, son linealmente independientes.
13. Pruebe que los $n + 1$ vectores de P_n: $1, x + a, (x + a), \ldots, (x + a)^n$, son linealmente independientes.
14. Demuestre que cualesquiera cuatro polinomios en P_2 son linealmente dependientes.
15. Compruebe que cualesquiera cinco matrices en $M_{2\times 2}$ son linealmente dependientes.
16. Sea V un espacio vectorial y $v_1, v_2, \ldots, v_n$ n vectores de V linealmente independientes. Demuestre que los vectores

$$u_1 = v_1$$
$$u_2 = v_1 + v_2$$
$$\vdots$$
$$v_n = v_1 + v_2 + \ldots v_n$$

son también linealmente independientes.

② 17. (El wronskiano). En este problema se desarrollará una herramienta especial para detectar la independencia lineal de un conjunto de k funciones en el espacio vectorial $C^n(\mathbf{R})$ ($k \leq n + 1$) —véase el ejercicio 5 de la sección anterior. Dadas k funciones $f_1, f_2, \ldots, f_k$ en $C^n(\mathbf{R})$, defina el *wronskiano* de estas funciones en $x \in \mathbf{R}$, denotado por $W(f_1, f_2, \ldots, f_k)(x)$, como el determinante

$$W(f_1, f_2, \ldots, f_k)(x) = \det \begin{bmatrix} f_1(x) & f_2(x) & \ldots & f_k(x) \\ f_1'(x) & f_2'(x) & \ldots & f_k'(x) \\ & \ldots & & \\ f_1^{(k-1)}(x) & f_2^{(k-1)}(x) & \ldots & f_k^{(k-1)}(x) \end{bmatrix}$$

Por ejemplo si $f_1(x) = e^{2x}$ y $f_2(x) = e^{3x}$ se tiene

$$W(e^{2x}, e^{3x})(x) = det \begin{bmatrix} e^{2x} & e^{3x} \\ 2e^{2x} & 3e^{3x} \end{bmatrix} = e^{5x}$$

a) Demuestre que si las k funciones $f_1, f_2, \ldots, f_k$ son linealmente dependientes entonces $W(f_1, f_2, \ldots, f_k)(x) = 0$ para toda $x \in \mathbf{R}$.
(*Sugerencia*: Considere la combinación lineal

$$c_1 f_1(x) + c_2 f_2(x) + \ldots + c_k f_k(x) = 0$$

Derive $k - 1$ veces esta expresión y obtenga así k expresiones del tipo

$$c_1 f_1^{(j)}(x) + c_2 f_2^{(j)}(x) + \ldots + c_k f_k^{(j)}(x) = 0 \qquad 0 \le j \le k - 1$$

Vea estas k expresiones como un sistema de k ecuaciones lineales en las incógnitas $c_1, c_2, \ldots, c_k$. Observe que el determinante del sistema es precisamente $W(f_1, f_2, \ldots, f_k)(x)$. Continúe . . .)

b) Compruebe que las cuatro funciones $f_1(x) = x^2 + 3x + 1$, $f_2(x) = 5x^2 - x + 3$, $f_3(x) = x^2 + 3x - 2$, $f_4(x) = 8x^2 + 3x - 1$, son linealmente dependientes (véase ejercicio 14). Calcule $W(f_1, f_2, f_3, f_4)(x)$.

c) Demuestre que la afirmación recíproca del inciso a) es falsa considerando el siguiente ejemplo: sean $f_1(x)$ y $f_2(x)$ las funciones siguientes:

$$f_1(x) = \begin{cases} x^2 & \text{si } x \ge 0 \\ 0 & \text{si } x < 0 \end{cases} \qquad f_2(x) = \begin{cases} 0 & \text{si } x > 0 \\ x_2 & \text{si } x \le 0 \end{cases}$$

c.1) Compruebe que $W(f_1, f_2)(x) = 0\ \forall\ x \in \mathbf{R}$.

c.2) Al usar directamente la definición de independencia lineal, demuestre que $f_1(x)$ y $f_2(x)$ son linealmente independientes.

d) Al usar el resultado del inciso a), demuestre que las siguientes funciones son linealmente independientes

d.1) $f_1(x) = e^{ax}, f_2(x) = e^{bx}$, con $a \ne b$

d.2) $f_1(x) = e^x, f_2(x) = xe^x, f_3(x) = x^2e^x$

d.3) $f_1(x) = \text{sen } x, f_2(x) = \cos x$

d.4) $f_1(x) = e^x \text{ sen } x, f_2(x) = e^x \cos x$

e) Al usar el wronskiano, resuelva nuevamente los ejercicios 12 y 13.

② 18. Demuestre la siguiente versión mejorada del teorema 4.1: considere el conjunto de n vectores ($n \ge 2$) en el espacio vectorial V $\{v_1, v_2, \ldots, v_n\}$. Suponga que $v_1 \ne 0$. Este conjunto es linealmente dependiente si, y sólo si existe un vector en él $v_j, j \ge 2$ tal que v_j es una combinación lineal de los vectores *precedentes* $v_1, v_2, \ldots, v_{j-1}$.

5. BASES Y DIMENSIÓN

DEFINICIÓN 5.1 Sea V un espacio vectorial. Se dice que el subconjunto S de V es una *base* de V si

1) S genera a V.
2) S es linealmente independiente.

Para ilustrar esta definición, vea algunos ejemplos de bases de espacios vectoriales.

EJEMPLO 1 Los vectores $v_1, = (3, 1, -1)$, $v_2 = (4, 1, 1)$ y $v_3 = (1, 2, 3)$ constituyen una base para el espacio vectorial $\mathbf{R}^3$.

Vea primeramente que estos vectores generan $\mathbf{R}^3$ (esto es, que $\mathcal{L}(v_1, v_2, v_3) = \mathbf{R}^3$).

Debe entonces mostrar que *todo* vector $(x, y, z) \in \mathbf{R}^3$ se puede escribir como combinación lineal de los vectores v_1, v_2, v_3, o sea, que dado $(x, y, z) \in \mathbf{R}^3$ existen escalares c_1, c_2, c_3 tales que

$$(x, y, z) = c_1(3, 1, -1) + c_2(4, 1, 1) + c_3(1, 2, 3)$$

Al realizar las operaciones indicadas e igualando las correspondientes coordenadas de los vectores involucrados en esta igualdad se obtiene

$$3c_1 + 4c_2 + c_3 = x$$
$$c_1 + c_2 + 2c_3 = y$$
$$-c_1 + c_2 + 3c_3 = z$$

Éste es un sistema no homogéneo de ecuaciones lineales en las incógnitas c_1, c_2 y c_3. Debe mostrar que independientemente de los valores de x, y, z, el sistema tiene solución. Sería suficiente entonces probar que el determinante de la matriz del sistema es diferente de cero (teorema 4.10 capítulo 1 y corolario 1 del teorema 2.7 capítulo 2).

En efecto

$$det \begin{bmatrix} 3 & 4 & 1 \\ 1 & 1 & 2 \\ -1 & 1 & 3 \end{bmatrix} = -15 \ne 0$$

Por lo tanto, los vectores v_1, v_2, v_3 generan $\mathbf{R}^3$.

Además, por el teorema 4.3 estos vectores son también linealmente independientes.

Entonces v_1, v_2 y v_3 constituyen una base de $\mathbf{R}^3$.

EJEMPLO 2 En $\mathbf{R}^n$, los vectores

$$e_1 = (1, 0, 0, \ldots, 0, 0)$$
$$e_2 = (0, 1, 0, \ldots, 0, 0)$$
$$\vdots$$
$$e_n = (0, 0, 0, \ldots, 0, 1)$$

forman una base, llamada *base canónica* (o estándar) de $\mathbf{R}^n$.

En efecto, dado $(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$ se puede escribir

$$(x_1, x_2, \ldots, x_n) = x_1e_1 + x_2e_2 + \ldots + x_ne_n$$

lo que muestra que $\mathcal{L}(e_1, e_2, \ldots, e_n) = \mathbf{R}^n$.

Por otra parte, la matriz cuya j-ésima columna es e_j $(j = 1, 2, \ldots, n)$ es la matriz identidad I_n, de modo que $\det I_n = 1 \neq 0$, y por el teorema 4.3, se concluye que los vectores $e_1, e_2, \ldots, e_n$ son linealmente independientes. Son, por tanto, una base de $\mathbf{R}^n$. Observe que en $\mathbf{R}^3$ los vectores de la base canónica son los conocidos vectores $e_1 = i = (1, 0, 0)$, $e_2 = j = (0, 1, 0)$, $e_3 = k = (0, 0, 1)$ que usaba en los cursos de física.

EJEMPLO 3 En el espacio $M_{m\times n}$, considere las $m \times n$ matrices siguientes: la matriz A_{ij} $(i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$ es la matriz de orden $m \times n$ que tiene todos sus elementos iguales a cero, excepto el que se encuentra en su i-ésima línea y j-ésima columna, que es igual a 1.

Es fácil verificar que estas $m \times n$ matrices forman una base de $M_{m\times n}$, llamada también base canónica de $M_{m\times n}$: dada la matriz $A = (a_{ij})_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}$ se puede escribir

$$A = \sum_{i,j=1}^{m,n} a_{ij}A_{ij}, \quad \text{lo que muestra que } \mathcal{L}(\{A_{ij},\ i = 1, \ldots, m; j = 1, \ldots, n\}) = M_{m\times n}.$$

Además la expresión $\sum_{i,j=1}^{m,n} c_{ij}A_{ij} = 0$ implica que $c_{ij} = 0$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, pues la matriz del lado izquierdo de esta expresión es la matriz $C = (c_{ij})_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}$

Esto muestra entonces que las matrices A_{ij}, $i = 1, \ldots, m, j = 1, \ldots, n$ son linealmente independientes.

Por ejemplo, en $M_{2\times 3}$, las $2 \times 3 = 6$ matrices que constituyen la base canónica son

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

EJEMPLO 4 Considere ahora los vectores $v_1 = 1$, $v_2 = x, \ldots, v_{n+1} = x^n$ en el espacio P_n. Se verá que constituyen una base de él.

En efecto, en la sección anterior se mostró que estos vectores son linealmente independientes. Además, dado cualquier polinomio $a_0 + a_1x + \ldots + a_nx^n$ en P_n es claro que

$$a_0 + a_1x + \ldots + a_nx^n = a_0v_1 + a_1v_2 + \ldots + a_nv_{n+1}$$

lo que muestra que $\mathcal{L}(1, x, \ldots, x^n) = P_n$.

Esta base también es llamada *base canónica* de P_n.

Vea un último ejemplo:

EJEMPLO 5 Considere el sistema homogéneo de ecuaciones lineales

$$x_1 + 3x_2 - 2x_3 + x_4 = 0$$
$$x_1 + x_2 - 4x_3 + 3x_4 = 0$$
$$-2x_1 - 8x_2 + 2x_3 = 0$$
$$-x_2 - x_3 + x_4 = 0$$

Se sabe que su conjunto solución forma un subespacio de $\mathbf{R}^4$ (véase sección 3). Se trata de encontrar una base para tal subespacio.

Primeramente, procédase a resolver el sistema usando el método de eliminación Gaussiana.

Se tiene

$$\begin{bmatrix} 1 & 3 & -2 & 1 \\ 1 & 1 & -4 & 3 \\ -2 & -8 & 2 & 0 \\ 0 & -1 & -1 & 1 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -5 & 4 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

de modo que las soluciones del sistema quedan descritas por la matriz $n \times 1$.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5r - 4t \\ -r + t \\ r \\ t \end{bmatrix} \quad r, t \in \mathbf{R}$$

o bien, identificando esta matriz con un vector de $\mathbf{R}^4$, se tiene

$$(5r - 4t, -r + t, r, t), \qquad r, t \in \mathbf{R}$$

Obsérvese que se puede escribir el vector solución como

$$(5r - 4t, -r + t, r, t) = r(5, -1, 1, 0) + t(-4, 1, 0, 1)$$

Llámese $v_1 = (5, -1, 1, 0)$ y $v_2 = (-4, 1, 0, 1)$.

Véase entonces que toda solución del sistema se puede escribir como combinación lineal de los vectores v_1 y v_2. Por lo tanto, estos vectores generan el espacio solución. Además es fácil verificar que ellos son linealmente independientes.

Entonces una base para el espacio solución del sistema está constituida por los vectores v_1 y v_2.

DEFINICIÓN 5.2 Si el espacio vectorial V posee una base formada por un número finito de vectores, se dice que V es un *espacio de dimensión finita*. Caso contrario se dice que V es un *espacio de dimensión infinita*.

Los ejemplos anteriores muestran entonces que $\mathbf{R}^n$, $M_{m\times n}$ y P_n son espacios de dimensión finita.

En este libro se centrará el estudio en espacios vectoriales de dimensión finita. Sin embargo, se puede ver ahora un ejemplo de un espacio vectorial de dimensión infinita.

Considérese el espacio P formado por *todos* los polinomios con coeficientes reales. Se verifica fácilmente que P es un espacio vectorial (del cual, para cada n, P_n es un subespacio de él). Se afirma que este espacio no posee una base con un número finito de vectores. Para ver la validez de esta afirmación supóngase por contradicción que tal base existe, dígase que es $\beta = \{v_1, v_2, \ldots, v_k\}$. En particular, esto significa que los vectores $v_1, v_2, \ldots, v_k$ generan al espacio P. Sea r un número mayor que el grado máximo de los polinomios $v_1, v_2, \ldots, v_k$, y sea p_r un polinomio de grado r con coeficientes reales. Entonces $p_r \in P$ y $p_r \notin \mathcal{L}(v_1, v_2, \ldots, v_k)$, lo que contradice el hecho de que β es una base de P (en particular contradice que los vectores $v_1, v_2, \ldots, v_k$ generan a P).

Es fácil ver que este espacio de polinomios tiene como base el conjunto *infinito*

$$S = \{1, x, x^2, \ldots, x^n, \ldots\}$$

Claramente $\mathcal{L}(S) = P$. Además, al usar el mismo argumento que se utilizó en la sección anterior para demostrar que $1, x, \ldots, x^n$ eran vectores de P_n linealmente independientes, se puede verificar que cualquier subconjunto finito de S es linealmente independiente. Por lo tanto (véase nota al final de la sección anterior) S es un conjunto linealmente independiente, que genera a P. Es entonces una base de P.

Considérese el espacio vectorial V y dígase que $\beta = \{v_1, v_2, \ldots, v_n\}$ es una base de él. En particular se tiene que $\mathcal{L}(v_1, v_2, \ldots, v_n) = V$. Esto significa que cada vector $v \in V$ se puede representar como una combinación lineal de los vectores $v_1, v_2, \ldots, v_n$. El primero de una serie de resultados importantes (e interesantes, por supuesto) que se verá a lo largo de esta sección, dice que esta representación de v como combinación lineal de los vectores de la base β es única. Aún más, se verá que esta unicidad en la representación de v caracteriza al hecho de que β es una base de V.

TEOREMA 5.1 El subconjunto $\beta = \{v_1, v_2, \ldots, v_n\}$ del espacio vectorial V es una base de V si, y sólo si cada vector $v \in V$ es expresado de manera única como una combinación lineal de los vectores de β.

DEMOSTRACIÓN Supóngase que $\beta = \{v_1, v_2, \ldots, v_n\}$ es una base de V. Sea $v \in V$ un vector arbitrario de V. Existen entonces escalares $c_1, c_2, \ldots, c_n$ tales que

$$v = c_1v_1 + c_2v_2 + \ldots + c_nv_n$$

Supóngase por contradicción, que v también tiene la representación

$$v = c_1'v_1 + c_2'v_2 + \ldots + c_n'v_n$$

Entonces

$$c_1v_1 + c_2v_2 + \ldots + c_nv_n = c_1'v_1 + c_2'v_2 + \ldots + c_n'v_n$$

o sea

$$(c_1 - c_1')v_1 + (c_2 - c_2')v_2 + \ldots + (c_n - c_n')v_n = 0 \qquad (5.1)$$

Hasta ahora sólo se ha usado el hecho de que $\mathcal{L}(v_1, v_2, \ldots, v_n) = V$. Pero siendo $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V, también se sabe que los vectores $v_1, v_2, \ldots, v_n$ son linealmente independientes. Entonces (5.1) implica que $c_i - c_i' = 0$, es decir $c_i = c_i'$, $i = 1, 2, \ldots, n$, lo que muestra entonces que la representación de $v \in V$ como combinación lineal de los vectores de la base β es única.

Recíprocamente, supóngase que cada $v \in V$ es representado de manera única como combinación lineal de $v_1, v_2, \ldots, v_n$. En tal caso se tiene obviamente que $\mathcal{L}(v_1, v_2, \ldots, v_n) = V$. Sólo falta verificar, entonces, que los vectores de β son linealmente independientes.

Considérese la combinación lineal

$$c_1v_1 + c_2v_2 + \ldots + c_nv_n = 0$$

Esta expresión se puede ver como la representación del vector $\vec{0} \in V$ como combinación lineal de los vectores $v_1, v_2, \ldots, v_n$. Pero obviamente el vector cero tiene también la representación

$$0 = 0 \cdot v_1 + 0 \cdot v_2 + \ldots + 0 \cdot v_n$$

En vista de la unicidad de la representación supuesta en la hipótesis, se concluye que $c_1 = c_2 = \ldots = c_n = 0$, es decir, que $v_1, v_2, \ldots, v_n$ son linealmente independientes y que por tanto constituyen una base de V.

Q.E.D.

El siguiente teorema, que relaciona el número de vectores que generan a un espacio vectorial V con el número máximo de vectores linealmente independientes que pueden existir en él, prepara el terreno para definir (como consecuencia del corolario que de él se deducirá) uno de los conceptos más importantes de este capítulo (y de todo el libro): El concepto de **dimensión**.

TEOREMA 5.2

Suponga que en el espacio vectorial V existen n vectores $v_1, v_2, \ldots, v_n$ que lo generan. Entonces cualquier conjunto linealmente independiente en V es finito y no contiene más de n vectores.

DEMOSTRACIÓN Considérese el conjunto de m vectores en V, $S = \{u_1, u_2, \ldots, u_m\}$, en donde $m > n$. Se mostrará entonces que S es linealmente dependiente.

Como $\mathcal{L}(v_1, v_2, \ldots, v_n) = V$, para cada u_j existen escalares $a_{ij, i=1,\ldots,n}$ tales que

$$u_j = a_{1j}v_1 + a_{2j}v_2 + \ldots + a_{nj}v_n = \sum_{i=1}^{n} a_{ij}v_i \quad j = 1, \ldots, m \tag{5.2}$$

Fórmese la combinación lineal

$$c_1u_1 + c_2u_2 + \ldots + c_mu_m \tag{5.3}$$

Al sustituir (5.2) en (5.3) se obtiene

$$c_1u_1 + c_2u_2 + \ldots + c_mu_m = \sum_{j=1}^{m} c_ju_j = \sum_{j=1}^{m} c_j \sum_{i=1}^{n} a_{ij}v_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} a_{ij}c_j\right)v_i \tag{5.4}$$

Sea A la matriz de orden $n \times m$ de coeficientes a_{ij}.

Considérese el sistema homogéneo de n ecuaciones lineales con m incógnitas $AX = 0$. Como se está suponiendo que $m > n$, existen entonces $c_1, c_2, \ldots, c_m$ *no todos cero* tales que $A\tilde{X} = 0$, donde $\tilde{X}$ es la matriz de orden $m \times 1$ de elementos $c_1, c_2, \ldots, c_m$.

Entonces se satisfacen los relaciones

$$\sum_{j=1}^{m} a_{ij}c_j = 0, \quad i = 1, 2, \ldots, n$$

Al sustituir estas expresiones en (5.4) se obtiene finalmente que $c_1u_1 + c_2u_2 + \ldots + c_mu_m = 0$, en donde no todos los escalares $c_1, c_2, \ldots, c_m$ son iguales a cero. Es decir, el conjunto S es linealmente dependiente.

Q.E.D.

COROLARIO

Sea V un espacio vectorial de dimensión finita. Entonces cualesquiera dos bases β_1 y β_2 de V, tienen el mismo número de vectores.

DEMOSTRACIÓN Sean $\beta_1 = \{v_1, v_2, \ldots, v_n\}$ y $\beta_2 = (u_1, u_2, \ldots, u_m)$ dos bases de V. Por una parte se tiene que $\mathcal{L}(\beta_1)$, $= V$ y β_2 es un conjunto linealmente independiente de vectores de V. Por el teorema anterior se tiene entonces $m \leq n$. Por otra parte $\mathcal{L}(\beta_2) = V$ y β_1 es un conjunto independiente de vectores de V. Nuevamente el teorema anterior permite concluir $n \leq m$. Entonces $n = m$.

Q.E.D.

Como ya se había dicho, este corolario permite hacer la siguiente importante definición:

DEFINICIÓN 5.3 Sea V un espacio vectorial de dimensión finita. Se define la *dimensión* de V como el número de vectores en alguna (y por tanto en cualquiera) base de V.

NOTACIÓN Se escribirá *dim* V para denotar la dimensión del espacio vectorial de dimensión finita V.

EJEMPLO 6

Por ejemplo, se había visto al inicio de la sección que los vectores $v_1 = (3, 1, -1)$, $v_2 = (4, 1, 1)$ y $v_3 = (1, 2, 3)$ en $\mathbf{R}^3$ formaban una base para este espacio. También se vio que $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ y $e_3 = (0, 0, 1)$ eran vectores de una base (la canónica) de $\mathbf{R}^3$. En ambos casos se tiene, pues, 3 vectores formando una base de $\mathbf{R}^3$ (tal como lo asegura el corolario anterior) y entonces $\dim \mathbf{R}^3 = 3$.

En general $\dim \mathbf{R}^n = n$, pues la base canónica de $\mathbf{R}^n$ contiene n vectores.

También se tiene que $\dim M_{m\times n} = m \times n$ y $\dim P_n = n + 1$, pues las bases canónicas de estos espacios contienen $m \times n$ y $n + 1$ vectores respectivamente.

NOTA: La **dimensión** del espacio trivial $\{0\}$ se *define* como siendo cero.

En el siguiente teorema se establecerán conexiones entre los conceptos "conjunto de generadores", "conjunto linealmente independiente" y "base" para un espacio vectorial de dimensión finita. De él se puede deducir un resultado "práctico" que facilitará la tarea de verificar si un conjunto dado de vectores en un espacio vectorial, es una base de él o no, cuando se sabe de antemano la dimensión del espacio. Pero antes de enunciarlo, se verán un par de lemas técnicos que se usarán en su demostración.

LEMA 1

Sea $S = \{v_1, v_2, \ldots, v_n\}$ un conjunto de generadores del espacio vectorial V, supóngase que existe v_j, $1 \leq j \leq n$ tal que puede escribirse como combinación lineal de los $n - 1$ vectores restantes. Entonces $\mathcal{L}(S\setminus\{v_j\}) = \mathcal{L}(S)$ (donde $S\setminus\{v_j\} = \{v \in S \mid v \neq v_j\}$).

DEMOSTRACIÓN La contención $\mathcal{L}(S\setminus\{v_j\}) \subset \mathcal{L}(S)$ es clara. Se verá entonces que $\mathcal{L}(S) \subset \mathcal{L}(S\setminus\{v_j\})$. Sea $v \in \mathcal{L}(S)$.

Existen escalares $c_1, c_2, \ldots, c_n$ tales que

$$v = c_1v_1 + c_2v_2 + \ldots + c_nv_n \tag{5.5}$$

Según la hipótesis del teorema, existen también escalares d_i, $i = 1, 2, \ldots n$, $i \neq j$ tales que

$$v_j = d_1v_1 + \ldots + d_{j-1}v_{j-1} + d_{j+1}v_{j+1} + \ldots + d_nv_n$$

Al sustituir esta expresión en (5.5) y agrupando, se obtiene al vector v expresado como combinación lineal de los $n - 1$ vectores $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n$. Entonces $v \in \mathcal{L}(S\backslash\{v_j\})$ como se quería demostrar.

Q.E.D.

LEMA 2

Sea S un subconjunto linealmente independiente del espacio vectorial V. Supóngase que el vector $v \in V$ no pertenece al espacio generado por S (esto es, $v \notin \mathcal{L}(S)$), entonces el conjunto $S' = S \cup \{v\}$ es linealmente independiente.

DEMOSTRACIÓN Sean $u_1, u_2, \ldots, u_k$ vectores de S. Considérese la combinación lineal

$$c_1u_1 + c_2u_2 + \ldots + c_ku_k + c_{k+1}v = 0 \tag{5.6}$$

Se afirma que $c_{k+1} = 0$. En efecto, si $c_{k+1} \neq 0$ se podría escribir

$$v = \left(-\frac{c_1}{c_{k+1}}\right)u_1 + \left(-\frac{c_2}{c_{k+1}}\right)u_2 + \ldots + \left(-\frac{c_k}{c_{k+1}}\right)u_k$$

lo que contradice la hipótesis de que $v \notin \mathcal{L}(S)$.

Entonces la expresión (5.6) queda como

$$c_1u_1 + c_2u_2 + \ldots + c_ku_k = 0$$

Como S es linealmente independiente, se concluye que $c_1 = \ldots = c_k = 0$, lo que muestra entonces que S' es linealmente independiente.

Q.E.D.

TEOREMA 5.3

Sea V un espacio vectorial de dimensión finita.

1) Cualquier conjunto de generadores de V contiene una base para V.
2) Si $\{u_1, u_2, \ldots, u_m\}$ es un conjunto linealmente independiente de V, existen vectores $w_1, w_2, \ldots, w_{n-m}$ ($n = dim\ V$) tales que $\beta = \{u_1, u_2, \ldots, u_m, w_1, w_2, \ldots, w_{n-m}\}$ es una base de V.

DEMOSTRACIÓN

1) Sea $S = \{v_1, v_2, \ldots, v_k\}$ un conjunto de generadores de V. Si S es linealmente independiente, entonces S es una base de V. Caso contrario, según el teorema 4.1 existe un vector v_j, $1 \leq j \leq k$ que se puede escribir como combinación lineal de los $k - 1$ vectores restantes. Sea $S_1 = S\backslash\{v_j\}$. Por el lema 1, $\mathcal{L}(S_1) = V$. Si S_1 es linealmente independiente, S_1 es entonces la base requerida. Caso contrario repita el proceso anterior. En algún momento se obtendrá un conjunto S_i $(1 \leq i \leq n - 1)$ linealmente independiente y ése será la base procurada. (El peor de los casos se presentaría cuando se pueda llegar a S_{n-1}, que es un conjunto con un solo vector. En este caso S_1 es linealmente independiente, y será la base requerida.)
2) Sea $S = \{u_1, u_2, \ldots, u_m\}$. Escríbase $V_0 = \mathcal{L}(S)$. Si $V_0 = V$, entonces S es la base requerida (pues S es linealmente independiente y $\mathcal{L}(S) = V$). Caso contrario (o sea si $\mathcal{L}(S)$ es un subespacio *propio* de V), existe un vector $w_1 \in V$ tal que $w_1 \notin \mathcal{L}(S)$. Según el lema 2, el conjunto $S_1 = S \cup \{w_1\}$ es linealmente independiente. Escríbase $V_1 = \mathcal{L}(S_1)$. Si $V_1 = V$, S_1 es la base requerida β. Caso contrario existe $w_2 \in V$, $w_2 \notin \mathcal{L}(S_1)$, etc. Al continuar este proceso de adjunción de vectores de V al conjunto S, se llegará a lo más en $n - m$ etapas, a un conjunto linealmente independiente que genere V. Ésta será la base β procurada.

Q.E.D.

La parte 2 del teorema anterior, dice que "todo conjunto linealmente independiente en un espacio vectorial de dimensión finita puede ser completado hasta formar una base del espacio". La misma demostración de este hecho dice cómo se debe proceder para obtener tal base.

Véase un ejemplo que ilustra las dos partes del teorema 5.3.

EJEMPLO 7

Sea V el espacio generado por los vectores $v_1 = (1, 2, 3, -1)$, $v_2 = (-1, 1, 0, 4)$, $v_3 = (3, 4, 7, -5)$, y $v_4 = (2, 5, 7, -1)$. V es, pues, un subespacio de $\mathbf{R}^4$.

La parte 1 del teorema 5.3 asegura que del conjunto de generadores $S = \{v_1, v_2, v_3, v_4\}$ se puede extraer una base para V.

Proceda como sigue:

Escriba la combinación lineal

$$c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 = 0$$

o

$$c_1(1, 2, 3, -1) + c_2(-1, 1, 0, 4) + c_3(3, 4, 7, -5) + c_4(2, 5, 7, -1) = 0 \tag{5.7}$$

de donde se obtiene

$$c_1 - c_2 + 3c_3 + 2c_4 = 0$$
$$2c_1 + c_2 + 4c_3 + 5c_4 = 0$$
$$3c_1 + \quad + 7c_3 + 7c_4 = 0$$
$$-c_1 + 4c_2 - 5c_3 - c_4 = 0$$

Al proceder por el método de eliminación Gaussiana se obtiene

$$\begin{bmatrix} 1 & -1 & 3 & 2 \\ 2 & 1 & 4 & 5 \\ 3 & 0 & 7 & 7 \\ -1 & 4 & -5 & -1 \end{bmatrix} \cdots \begin{bmatrix} 1 & 0 & 7/3 & 7/3 \\ 0 & 1 & -2/3 & 1/3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

por lo que la solución del sistema es:

$$c_1 = -\frac{7}{3}r - \frac{7}{3}t$$

$$c_2 = \frac{2}{3}r - \frac{1}{3}t$$

$$c_3 = r \qquad\qquad r, t \in \mathbf{R}$$

$$c_4 = t$$

En particular se tienen dos conjuntos solución $\{c_1 = -7, c_2 = -1, c_3 = 0, c_4 = 3\}$ y $\{c_1 = -7, c_2 = 2, c_3 = 3, c_4 = 0\}$ que sustituidos en (5.7) dan

$$v_4 = \frac{7}{3}v_1 + \frac{1}{3}v_2$$

$$v_3 = \frac{7}{3}v_1 - \frac{2}{3}v_2$$

Entonces $\mathcal{L}(v_1, v_2, v_3, v_4) = \mathcal{L}(v_1, v_2)$ y $\beta = \{v_1, v_2\}$ es una base de V, como lo asegura la parte 1 del teorema 5.3. El conjunto $T = \{v_1, v_2\}$ es entonces linealmente independiente. La parte 2 del teorema 5.3 asegura que se puede extender T hasta completar una base β de $\mathbf{R}^4$. Se tendrá que añadir $dim\ \mathbf{R}^4 - 2 = 4 - 2 = 2$ vectores a T para completar la base β. Por simplicidad, escójase $w_1 = (1, 0, 0, 0)$. Sólo se tiene que verificar que $w_1 \notin \mathcal{L}(v_1, v_2)$, es decir, que no existen escalares c_1 y c_2 tales que $w_1 = c_1v_1 + c_2v_2$. Se escribe entonces

$$(1, 0, 0, 0) = c_1(1, 2, 3, -1) + c_2(-1, 1, 0, 4)$$

de donde se obtiene el sistema

$$\begin{aligned} c_1 - c_2 &= 1 \\ 2c_1 + c_2 &= 0 \\ 3c_1 \quad &= 0 \\ -c_1 + 4c_2 &= 0 \end{aligned}$$

que se comprueba fácilmente que no tiene solución.

Se tiene entonces el conjunto $T_1\{v_1, v_2, w_1\}$ que es (según el lema 2) linealmente independiente. Escójase ahora $w_2 = (0, 1, 0, 0)$, y verifíquese que $w_2 \notin \mathcal{L}(v_1, v_2, w_1)$. Al escribir

$$(0, 1, 0, 0) = c_1(1, 2, 3, -1) + c_2(-1, 1, 0, 4) + c_3(1, 0, 0, 0)$$

se obtiene el sistema

$$\begin{aligned} c_1 - c_2 + c_3 &= 0 \\ 2c_1 + c_2 &= 1 \\ 3c_1 \quad &= 0 \\ -c_1 + 4c_2 &= 0 \end{aligned}$$

para el cual se verifica también fácilmente que no tiene solución. Se tiene ahora entonces el conjunto linealmente independiente $T_2 = \{v_1, v_2, w_1, w_2\}$. Éste es un conjunto linealmente independiente que *genera* a $\mathbf{R}^4$ (caso contrario existiría $w_3 \in \mathbf{R}^4$ tal que $w_3 \notin \mathcal{L}(v_1, v_2, w_1, w_2)$. El lema 2 diría entonces que el conjunto $T_3 = \{v_1, v_2, w_1, w_2, w_3\}$ es linealmente independiente. Pero esto contradice el teorema 5.2, pues los vectores e_1, e_2, e_3, e_4 de la **base canónica** generan a $\mathbf{R}^4$ o si no, por el argumento del corolario que se establecerá enseguida). Del teorema 5.3 se deduce el siguiente importante corolario:

COROLARIO

Sea V un espacio vectorial de dimensión finita. Dígase que $dim\ V = n$. Entonces

1) Cualquier conjunto con n vectores linealmente independientes es una base de V.
2) Cualquier conjunto con n vectores que genera a V, es una base de V.

DEMOSTRACIÓN

1) Según el teorema anterior, todo conjunto linealmente independiente puede ser extendido para completar una base. Pero, por una parte, V tiene n vectores en cualquier base y, por otra, el conjunto linealmente independiente del que se parte tiene n vectores. Este debe ser entonces una base de V.
2) Por el teorema anterior, de cualquier conjunto de vectores que generan a V se puede extraer una base. Por el mismo argumento que en 1), el conjunto formado por n (= número de vectores en una base de V) vectores que generan a V debe ser una base de V.

Q.E.D.

Este corolario muestra entonces que si $dim\ V = n$, para verificar que un conjunto $S = \{v_1, \ldots, v_n\}$ es una base de V, sólo se tiene que mostrar que

1) S es un conjunto linealmente independiente o (y no "y" como en la definición 5.1).
2) S genera a V.

Al usar estas ideas, se puede dar una nueva versión mejorada del teorema 4.3 visto en la sección anterior.

TEOREMA 4.3

(VERSIÓN MEJORADA DEL TEOREMA 4.3.)
Sean $v_j = (a_{1j}, a_{2j}, \ldots, a_{nj})$ $j = 1, 2, \ldots, n$, n vectores en $\mathbf{R}^n$. Considérese la matriz A de orden n $A = (a_{ij})_{i,j=1,\ldots,n}$. Entonces, los vectores $v_1, v_2, \ldots, v_n$ *forman una base de* $\mathbf{R}^n$ si, y sólo si $det\, A \neq 0$.

DEMOSTRACIÓN Consecuencia inmediata del teorema 4.3 y de la parte 1 del corolario anterior.

Q.E.D.

EJEMPLO 8 Por ejemplo, los vectores $v_1 = (a, c)$ y v_2, (b, d) en $\mathbf{R}^2$ forman una base de este espacio si, y sólo si $ad - bc \neq 0$.

El siguiente teorema dice que si se tiene un espacio de dimensión finita V, con sus subespacios acontece lo que es natural esperar que acontezca: que sean también de dimensión finita y que su dimensión no sea mayor que la dimensión de V.

TEOREMA 5.4

Sea W un subespacio del espacio vectorial de dimensión finita V. Entonces W es de dimensión finita y $dim\, W \leq dim\, V$.

DEMOSTRACIÓN El caso $W = \{0\}$ es obvio, pues $dim\, V \geq 0$. La idea del argumento que se usará para mostrar el teorema es la misma que la usada en la demostración del teorema 5.3(2). Tómese $w_1 \in W$. Si $\mathcal{L}(w_1) = W$ entonces $\{w_1\}$ es una base de W (y por tanto W es de dimensión finita) que puede completarse para formar una base de V (teorema 5.3). En tal caso se tiene $dim\, W \leq dim\, V$. Si $\mathcal{L}(w_1) \neq W$, tómese $w_2 \in W$, $w_2 \notin \mathcal{L}(w_1)$. El conjunto $\{w_1, w_2\}$ es linealmente independiente (lema 2 anterior). Si $\mathcal{L}(w_1, w_2) = W$ entonces $\{w_1, w_2\}$ es una base de W (y entonces W es de dimensión finita) que se puede completar para formar una base de V (nuevamente teorema 5.3). En este caso, $dim\, W \leq dim\, V$. Continúese con este proceso tantas veces como sea necesario hasta obtener una base de W. Obsérvese que no se puede tener más de $dim\, V$ etapas. Esto muestra que W es de dimensión finita. Con la base obtenida de W, complétese hasta formar una base de V. Entonces $dim\, W \leq dim\, V$.

Q.E.D.

COROLARIO 1

Sea W un subespacio del espacio vectorial de dimensión finita V. Si $dim\, W = dim\, V$, entonces $W = V$.

DEMOSTRACIÓN Inmediata, del teorema anterior.

Q.E.D.

COROLARIO 2 El espacio $C(\mathbf{R})$ de todas las funciones reales continuas definidas en la recta, es de dimensión infinita.

DEMOSTRACIÓN Se ha visto ya que el espacio P de todos los polinomios con coeficientes reales es un espacio vectorial de dimensión infinita. Es claro que P es un subespacio de $C(\mathbf{R})$. Si $C(\mathbf{R})$ fuera de dimensión finita, el teorema anterior implicaría que P es de dimensión finita, lo cual, como se dijo, es falso. Se sigue entonces que $C(\mathbf{R})$ es de dimensión infinita.

Q.E.D.

Se puede también deducir de este teorema, el siguiente resultado geométrico que ya se había prometido en la sección 3.

COROLARIO 3

Los únicos subespacios no triviales de $\mathbf{R}^3$ son rectas que pasan por el origen y planos que pasan por el origen.

DEMOSTRACIÓN Sea W un subespacio de $\mathbf{R}^3$. Según el teorema 5.4 $dim\, W$ es igual a 0, 1, 2 o 3. Si $dim\, W = 0$, entonces $W = \{0\}$. Si $dim\, W = 3$, entonces $W = \mathbf{R}^3$ (¿por qué?). Por tanto, si W es un subespacio no trivial de $\mathbf{R}^3$ su dimensión debe ser 1 o 2. Considérense estos dos casos por separado.

$dim\, W = 1$. En este caso, cualquier base de W contiene un solo vector $v = (a, b, c)$ en W (en particular, a, b y c no son simultáneamente cero). Entonces $\mathcal{L}(v) = W$. O sea que

$$W = \{(x, y, z) \mid (x, y, z) = t(a, b, c) \quad t \in \mathbf{R}\}$$
$$= \{(x, y, z) \mid x = ta,\ y = tb,\ z = tc \quad t \in \mathbf{R}\}$$

W es, pues, una recta que pasa por el origen.

$dim\, W = 2$. En este caso, cualquier base de W contiene dos vectores $v_1 = (a_1, b_1, c_1)$ y $v_2 = (a_2, b_2, c_2)$. Entonces $W = \mathcal{L}(v_1, v_2)$. O sea

$$W = \{(x, y, z) \mid (x, y, z) = \alpha_1(a_1, b_1, c_1) + \alpha_2(a_2, b_2, c_2),\ \alpha_1, \alpha_2 \in \mathbf{R}\}$$
$$= \{(x, y, z) \mid x = \alpha_1 a_1 + \alpha_2 a_2,\ y = \alpha_1 b_1 + \alpha_2 b_2, z = \alpha_1 c_1 + \alpha_2 c_2,\ \alpha_1, \alpha_2 \in \mathbf{R}\}$$

Se puede describir más explícitamente este subespacio si se considera el sistema

$$\alpha_1 a_1 + \alpha_2 a_2 = x$$
$$\alpha_1 b_1 + \alpha_2 b_2 = y$$
$$\alpha_1 c_1 + \alpha_2 c_2 = z$$

¿Qué condiciones deben cumplir x, y, z para que este sistema tenga solución para α_1 y α_2? (esto es, ¿cómo debe ser el vector $(x, y, z) \in \mathbf{R}^3$ para que $(x, y, z) \in \mathscr{L}(v_1, v_2)$?). Al usar el método de eliminación Gaussiana se puede ver que tal condición es que

$$(b_1c_2 - b_2c_1)x + (c_1a_2 - c_2a_1)y + (b_2a_1 - b_1a_2)z = 0$$

Llámese $A = b_1c_2 - b_2c_1$, $B = c_1a_2 - c_2a_1$, $C = b_2a_1 - b_1a_2$. Obsérvese que el hecho de que v_1, v_2 forman una *base* de W, asegura que A, B y C no pueden ser simultáneamente cero (¿por qué?). Entonces W queda descrito en este caso como $W\{x, y, z\} \mid Ax + By + Cz = 0\}$ que es un plano que pasa por el origen.

Q.E.D.

NOTA: Con un argumento similar podría ser redemostrado el teorema 3.2, de manera completamente distinta a como se hizo en la sección 3. Se deja al lector que lo haga como ejercicio.

En el siguiente teorema se quieren resaltar los resultados más importantes que han aparecido en esta sección. Su demostración está "dispersa" en las páginas anteriores.

TEOREMA 5.5

Sea V un espacio vectorial de dimensión finita. Dígase que $dim\ V = n$. Entonces:

1. Cualquier base de V tiene n vectores.
2. Cualquier conjunto linealmente independiente con n vectores es una base de V.
3. Cualquier conjunto con n vectores que genere a V, es una base de él.
4. No existen conjuntos linealmente independientes en V con más de n vectores.
5. No existen conjuntos con menos de n vectores que generen a V.
6. Cualquier subconjunto de V linealmente independiente, es subconjunto de una base de V.
7. Cualquier conjunto que genere a V, tiene como subconjunto una base de V.
8. Cualquier subespacio de V es de dimensión finita y su dimensión es no mayor que n.

DEMOSTRACIÓN Revise el material de esta sección.

Q.E.D.

Sea V un espacio vectorial de dimensión finita. Sean W_1 y W_2 subespacios de él. Como se había visto en la sección anterior, $W_1 + W_2$ y $W_1 \cap W_2$ son también subespacios de V. Por el corolario 1 del teorema 5.4, los subespacios W_1, W_2, $W_1 + W_2$ y $W_1 \cap W_2$ son todos de dimensión finita. El siguiente teorema dice cómo están relacionadas sus dimensiones:

TEOREMA 5.6

Sean W_1 y W_2 subespacios del espacio vectorial V de dimensión finita. Entonces

$$dim\,(W_1 + W_2) = dim\ W_1 + dim\ W_2 - dim\ W_1 \cap W_2$$

DEMOSTRACIÓN Sea $r = dim\,(W_1 \cap W_2)$. Tómese una base cualquiera de $W_1 \cap W_2$, dígase $\beta_0 = \{v_1, v_2, \ldots, v_r\}$. Ya se había visto que $W_1 \cap W_2$ puede ser contemplado como un subespacio de W_1 y de W_2. Por el teorema 5.3 se puede entonces extender β_0 para obtener una base β_1 de W_1 y β_2 de W_2. Dígase que tales bases son $\beta_1 = \{v_1, \ldots, v_r, x_1, \ldots, x_n\}$ y $\beta_2 = \{v_1, \ldots, v_r, y_1, \ldots, y_m\}$. Obsérvese entonces que $dim\ W_1 = r + n$ y $dim\ W_2 = r + m$. Se mostrará entonces que el conjunto $\beta = \beta_0 \cup \beta_1 \cup \beta_2$ es una base de $W_1 + W_2$. Véase primeramente que los vectores de β generan a $W_1 + W_2$. Tómese $v \in W_1 + W_2$. Entonces $v = w_1 + w_2$, $w_1 \in W_1$, $w_2 \in W_2$. Se puede escribir w_1 y w_2 en términos de las bases β_1 y β_2 respectivamente como $w_1 = c_1v_1 + \ldots + c_rv_r + d_1x_1 + \ldots + d_nx_n$ y $w_2 = c_1'v_1 + \ldots + c_r'v_r + d_1'y_1 + \ldots + d_m'y_m$, de modo que $v = w_1 + w_2 = (c_1 + c_1')v_1 + \ldots + (c_r + c_r')v_r + d_1x_1 + \ldots + d_nx_n + d_1'y_1 + \ldots + d_m'y_m$ lo que muestra que $\mathscr{L}(\beta) = W_1 + W_2$. Se probará ahora que los vectores de β son linealmente independientes. Tómese la combinación lineal de los vectores de $a_1v_1 + \ldots + a_rv_r + b_1x_1 + \ldots + b_nx_n + c_1y_1 + \ldots + c_my_m = 0$. Reescríbase esta expresión, como

$$-c_1y_1 - \ldots - c_my_m = a_1v_1 + \ldots + a_rv_r + b_1x_1 + \ldots + b_nx_n \qquad (5.8)$$

Ciertamente el vector del miembro izquierdo de esta expresión pertenece a W_2. Pero también esta expresión dice que tal vector se escribe como una combinación lineal de los vectores de la base β_1 de W_1. Por lo tanto

$$-c_1y_1 - \ldots - c_my_m \in W_1 \cap W_2$$

y entonces, como β_0 es una base de $W_1 \cap W_2$ existen escalares $k_1, \ldots, k_r$ tales que

$$-c_1y_1 - \ldots - c_my_m = k_1v_1 + \ldots + k_rv_r$$

o bien,

$$k_1v_1 + \ldots + k_rv_r + c_1y_1 + \ldots + c_my_m = 0$$

Obsérvese que ésta es una combinación lineal de los vectores de β_2 igualada a cero. Tales vectores son, en particular, linealmente independientes. Entonces se concluye que $k_1 = \ldots = k_r = c_1 = \ldots = c_m = 0$. De la expresión (5.8) se puede escribir $a_1v_1 + \ldots + a_rv_r + b_1x_1 + \ldots + b_nx_n = 0$ y como los vectores de β_1 son linealmente independientes, se concluye finalmente que $a_1 = \ldots = a_r = b_1 = \ldots = b_n = (c_1$

$= \ldots = c_m) = 0$. Es decir, que los vectores de β son linealmente independientes. Entonces, β es una base de $W_1 + W_2$ y por lo tanto

$$\begin{aligned} dim\,(W_1 + W_2) &= r + n + m \\ &= (r + n) + (r + m) - r \\ &= dim\,W_1 + dim\,W_2 - dim\,W_1 \cap W_2 \end{aligned}$$

Q.E.D.

EJEMPLO 9

Sea por ejemplo $V = M_{2\times 2}$. Sea W_1 el subespacio de las matrices triangulares superiores y W_2 el subespacio de las matrices triangulares inferiores. En la sección anterior se vio que $W_1 \cap W_2$ = subespacio de las matrices diagonales y que $W_1 + W_2 = M_{2\times 2}$. En este caso, se tiene que $dim(W_1 + W_2) = dim\,V = dim\,M_{2\times 2} = 4$.

La dimensión de W_1 es 3, pues una base para este subespacio es:

$$\left\{\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right\}$$

y la dimensión de W_2 es 3, pues una base de este subespacio es:

$$\left\{\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right\}$$

Similarmente, la dimensión de $W_1 \cap W_2$ es 2, pues las matrices

$$\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$$

constituyen una base para el subespacio de matrices diagonales. Entonces

$$\begin{array}{ccccccc} dim\,(W_1 + W_2) & = & dim\,W_1 & + & dim\,W_2 & - & dim\,W_1 \cap W_2 \\ \| & & \| & & \| & & \| \\ 4 & & 3 & & 3 & & 2 \end{array}$$

Un corolario inmediato del teorema 5.6 es el siguiente, cuya demostración se deja como un sencillo ejercicio para el lector.

COROLARIO

> Sean W_1 y W_2 subespacios del espacio vectorial de dimensión finita V tales que $W_1 \oplus W_2 = V$. Entonces $dim\,V = dim\,W_1 + dim\,W_2$.

EJERCICIOS (SECCIÓN 5, CAPÍTULO 3)

1. Determine si los siguientes conjuntos de vectores en $M_{2\times 2}$ constituyen una base de este espacio

a) $v_1 = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$ $v_2 = \begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$ $v_3 = \begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}$ $v_4 = \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$

b) $v_1 = \begin{bmatrix}2 & 1\\ 0 & 0\end{bmatrix}$ $v_2 = \begin{bmatrix}0 & 0\\ 3 & 1\end{bmatrix}$ $v_3 = \begin{bmatrix}1 & 0\\ 3 & 2\end{bmatrix}$ $v_4 = \begin{bmatrix}0 & -1\\ 0 & 3\end{bmatrix}$

c) $v_1 = \begin{bmatrix}2 & 1\\ 3 & 2\end{bmatrix}$ $v_2 = \begin{bmatrix}-1 & 3\\ 0 & 1\end{bmatrix}$ $v_3 = \begin{bmatrix}1 & -1\\ 1 & -1\end{bmatrix}$ $v_4 = \begin{bmatrix}2 & 3\\ -1 & 2\end{bmatrix}$

d) $v_1 = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ $v_2 = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ $v_3 = \begin{bmatrix}2 & 3\\ 3 & 2\end{bmatrix}$ $v_4 = \begin{bmatrix}-1 & 4\\ 5 & 3\end{bmatrix}$

2. a) Establezca una base del espacio S dado por

$$S = \{(x, y, z) \mid x + y + z = 0\}$$

b) Determine la dimensión del espacio W dado por

$$W = \{(x, y, z, w) \mid x + y + z + w = 0\}$$

3. Establezca una base del espacio solución de cada uno de los sistemas homogéneos de ecuaciones lineales siguientes:

a) $\begin{aligned} x_1 + x_2 &= 0 \\ 2x_1 + 2x_2 &= 0 \end{aligned}$

b) $\begin{aligned} 3x_1 + x_2 - x_3 &= 0 \\ x_1 - x_2 + x_3 &= 0 \end{aligned}$

c) $\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ 2x_1 - x_2 + 8x_3 &= 0 \\ 4x_1 - 3x_2 + 10x_3 &= 0 \end{aligned}$

d) $\begin{aligned} x_1 + 2x_2 + x_3 - 2x_4 &= 0 \\ 2x_1 - x_2 + x_3 + x_4 &= 0 \\ x_1 - 2x_2 + 2x_3 - x_4 &= 0 \end{aligned}$

e) $\begin{aligned} x_1 + x_2 - x_3 + 2x_4 &= 0 \\ 5x_1 + 4x_2 - 2x_3 - 2x_4 &= 0 \\ 2x_1 - x_2 + x_3 - 8x_4 &= 0 \\ 5x_1 - x_2 + x_3 - 14x_4 &= 0 \end{aligned}$

① 4. Demuestre que el conjunto vacío es una base del espacio vectorial trivial $\{0\}$. (*Sugerencia*: proceda por contradicción suponiendo que existe una base β de $\{0\}$ y que $\beta \neq \emptyset$ . . . [Véase también el ejercicio 29(c) de la sección 3.]

5. Sea A una matriz inversible de orden n. Encuentre una base para el espacio solución del sistema homogéneo de ecuaciones lineales $AX = 0$.

6. Para cada uno de los siguientes conjuntos de vectores en $\mathbb{R}^3$, existen razones "evidentes" por las cuales tal conjunto no puede constituir una base de este espacio. Diga cuáles son estas razones.

a) $\{(1, 3, 2), (2, 1, 1)\}$
b) $\{(1, 1, 1), (2, 2, 2), (3, 3, 3)\}$
c) $\{(2, 1, 3), (0, 0, 0), (3, 1, -1)\}$
d) $\{(2, 1, 4), (-8, 1, 3), (8, 0, 7), (3, 1, 1)\}$
e) $\{(3, 1, 4), (7, 4, 2), (3, 1, 4)\}$

7. En cada uno de los ejercicios siguientes encuentre una base β de $\mathbf{R}^3$ tal que β contenga al (a los) vector(es) indicado(s)

a) $v_1 = (1, 1, 1)$
b) $v_1 = (0, 1, 0)$, $v_2 = (1, 0, 1)$
c) $v_1 = (2, 3, 1)$, $v_2 = (3, 1, 2)$
d) $v_1 = (1, 1, 1)$, $v_2 = (3, -1, 4)$, $v_3 = (0, 1, 2)$

8. a) ¿Existe alguna base de $\mathbf{R}^3$ que contenga a los vectores $v_1 = (2, 1, 3)$, y $v_2 = (4, 2, 6)$? Explique.
b) ¿Existe alguna base de $\mathbf{R}^3$ que contenga a los vectores $v_1 = (3, 1, 2)$, $v_2 = (-1, 2, 0)$, $v_3 = (0, 3, 4)$ y $v_4 = (3, -2, 8)$? Explique.

9. Considere el siguiente conjunto de vectores en $\mathbf{R}^3$

$$S = \{(2, 1, 1), (10, 0, 2), (3, -1, 0), (-3, 6, 3), (0, 1, -1)\}$$

a) Demuestre que S es un conjunto de generadores de $\mathbf{R}^3$.
b) Extraiga una base para $\mathbf{R}^3$ del conjunto S.

10. Considere el subespacio W de $\mathbf{R}^4$ generado por los vectores $v_1 = (0, 1, -3, 2)$, $v_2 = (1, -1, 0, 1)$, $v_3 = (3, 0, 1, -1)$ y $v_4 = (4, 0, -2, 2)$.
a) Obtenga una base β de W formada por algunos de estos generadores.
b) Complete β para formar una base de $\mathbf{R}^4$.

11. Considere el subespacio W de $\mathbf{R}^5$ generado por los vectores $v_1 = (1, -1, 1, 0, 3)$, $v_2 = (0, 3, -2, -1, 4)$, $v_3 = (1, 11, -7, -4, 19)$ y $v_4 = (3, 0, 1, -1, 13)$.
a) Obtenga una base β de W formada por algunos de estos generadores.
b) Complete β para formar una base de $\mathbf{R}^5$.

12. Considere el subespacio W de $M_{2\times 3}$ generado por los vectores

$$v_1 = \begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & 4 \end{bmatrix} \quad v_3 = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -3 & 2 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} 1 & 4 & 4 \\ 5 & -2 & 7 \end{bmatrix} \quad v_4 = \begin{bmatrix} -1 & 3 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

$$v_5 = \begin{bmatrix} -4 & 3 & -9 \\ -6 & 5 & -11 \end{bmatrix}$$

a) Obtenga una base β de W formada por algunos de estos generadores.
b) Complete β para formar una base de $M_{2\times 3}$.

13. Considere el subespacio de P_n generado por los vectores $1, x + 1, (x + 1)^2, \ldots, (x + 1)^n$

a) Obtenga una base β de W formada por algunos de estos generadores.
b) Complete β para formar una base de P_n.
(*Sugerencia*: véase el ejercicio 13 de la sección anterior.)

14. Considere los subespacios W_1 y W_2 de $M_{2\times 2}$ dados por

$$W_1 = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid d = 0\right\} \qquad W_2 = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a = 0\right\}$$

a) Describa los subespacios $W_1 + W_2$ y $W_1 \cap W_2$.
b) Determine bases para cada uno de los subespacios W_1 y W_2, así como para los subespacios $W_1 + W_2$ y $W_1 \cap W_2$.
c) Verifique que se satisface el teorema 5.6.

15. Considere los subespacios W_1 y W_2 de $M_{3\times 3}$ dados por

$$W_1 = \mathcal{L}\left(\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)$$

$$W_2 = \mathcal{L}\left(\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)$$

a) Describa los subespacios $W_1 + W_2$ y $W_1 \cap W_2$.
b) Determine bases para cada uno de los subespacios W_1, W_2, $W_1 + W_2$ y $W_1 \cap W_2$.
c) Verifique que se satisface el teorema 5.6.

③ 16. Considere los subespacios W_1 y W_2 de $M_{n\times n}$ dados por

$$W_1 = \{A \in M_{n\times n} \mid A = A^t\} \quad \text{(matrices simétricas)}$$
$$W_2 = \{A \in M_{n\times n} \mid A = -A^t\} \quad \text{(matrices antisimétricas)}$$

a) Determine bases para W_1 y W_2.
b) Demuestre que

$$dim\, W_1 = \frac{n(n+1)}{2} \qquad dim\, W_2 = \frac{n(n-1)}{2}$$

c) Demuestre que

$$M_{n\times n} = W_1 \oplus W_2$$

(*Sugerencia*: Si A es cualquier matriz de orden n, la matriz $A + A^t$ es una matriz simétrica y $A - A^t$ es antisimétrica.)
d) Verifique que se satisface el teorema 5.6 (su corolario).

17. Sean P_1 y P_2 dos planos distintos en $\mathbf{R}^3$ que pasan por el origen. Demuestre que

$$\mathbf{R}^3 = P_1 + P_2$$

Usando el teorema 5.6, demuestre también que $P_1 \cap P_2$ es una recta que pasa por el origen.

② 18. ¿Cierto o falso? Sean W_1 y W_2 dos subespacios distintos del espacio vectorial V. Si $\beta_1 = \{v_1, v_2, v_3, w_1, w_2\}$ es una base de W_1 y $\beta_2 = \{v_1, v_2, v_3, u_1, u_2\}$ es una base de W_2, entonces $\beta = \{v_1, v_2, v_3\}$ es una base de $W_1 \cap W_2$.

19. ¿Cierto o falso? Sean W_1 y W_2 dos subespacios distintos del espacio vectorial V. Si β_1 es una base de W_1 y β_2 es una base de W_2, entonces $\beta = \beta_1 \cup \beta_2$ es una base de $W_1 + W_2 = \mathcal{L}(W_1 \cup W_2)$.

6. COORDENADAS

A partir de esta sección y por el resto del libro, se va a considerar a una base en un espacio vectorial V de dimensión finita no sólo como un conjunto de vectores del espacio, sino como un conjunto *ordenado* de vectores de V (es decir, un conjunto de vectores en el cual está establecido un orden entre ellos de modo que tenga sentido hablar del "primer vector", "segundo vector", etcétera).

Con esta perspectiva se tendrá, por ejemplo, que la base $\{e_1, e_2\}$ (la canónica) de $\mathbf{R}^2$ es *diferente* de la base $\{e_2, e_1\}$.

Considérese un espacio vectorial V de dimensión finita. Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V. Se ha visto que para cada $v \in V$ existen (únicos; teorema 5.1) escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ tales que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

Se puede establecer una identificación entre los vectores v del espacio V y los vectores de $\mathbf{R}^n$ cuya i-ésima coordenada es el escalar α_i que multiplica al vector v_i (el i-ésimo vector) en la expresión que define a v como combinación lineal de los elementos de la base $\beta = \{v_1, v_2, \ldots, v_n\}$. Es decir

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \in V \xleftrightarrow{\text{identificación}} (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbf{R}^n$$

Obsérvese primeramente que esta identificación se está haciendo *vía una base concreta de V*. Por tanto, si se cambia la base β al mismo vector $v \in V$ le corresponde un vector distinto de $\mathbf{R}^n$ (la identificación cambia).

Por otra parte, obsérvese también que esta identificación está bien definida, pues se ha dado previamente un orden a los vectores de la base β, de modo que cada vector $v \in V$ tendrá asociado (vía la base β) un vector bien determinado en $\mathbf{R}^n$ (una n-ada *ordenada* de números reales).

Por ejemplo, si el espacio V fuera de dimensión 2, y se toma $\beta = \{v_1, v_2\}$ una base de V, al vector $v = c_1 v_1 + c_2 v_2$ le corresponde, según la identificación mencionada, el vector $(c_1, c_2) \in \mathbf{R}^2$. Sin embargo, si se altera el orden de los vectores de β, se está (como se dijo anteriormente) considerando otra base *distinta* de V, dígase $\beta' = \{v_2, v_1\}$ respecto de la cual el vector v se identifica con $(c_2, c_1) \in \mathbf{R}^2$.

DEFINICIÓN 6.1 Sea V un espacio vectorial de dimensión n y sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V. Defínase el *vector de coordenadas de v respecto de la base* β, el cual se denota por $(v)_\beta$, como el vector

$$(v)_\beta = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbf{R}^n$$

en donde los escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ son tales que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

También se va a considerar la *matriz de coordenadas de v respecto de la base* β, denotada por $[v]_\beta$ como la matriz $n \times 1$

$$[v]_\beta = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Sin mayores explicaciones, se usará $(v)_\beta$ o $[v]_\beta$ según convenga en cada caso particular que se esté considerando.

EJEMPLO 1 Por ejemplo, sea $V = P_4$.

Se sabe que *dim* $V = 5$. Es fácil verificar que los vectores

$$v_1 = 1,\ v_2 = 1 + x,\ v_3 = 1 + x + x^2,\ v_4 = 1 + x + x^2 + x^3,\ v_5 = 1 + x + x^2 + x^3 + v^4$$

forman una base de P_4, la cual se denotará por $\beta = \{v_1, v_2, v_3, v_4, v_5\}$. Dado el vector de P_4

$$v = 2 + 3x + 2x^2 + 5x^3 - 2x^4$$

obtenga $(v)_\beta$.

Todo lo que se tiene que hacer es hallar los escalares $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ tales que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 + \alpha_5 v_5$$

Al escribir explícitamente esta expresión se obtiene

$$2 + 3x + 2x^2 + 5x^3 - 2x^4 = \alpha_1(1) + \alpha_2(1 + x) + \alpha_3(1 + x + x^2)$$
$$+ \alpha_1(1 + x + x^2 + x^3) + \alpha_5(1 + x + x^2 + x^3 + x^4)$$

$$= (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5) + (\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5)x + (\alpha_3 + \alpha_4 + \alpha_5)x^2 + (\alpha_4 + \alpha_5)x^3 + \alpha_5 x^4$$

de donde

$$\left.\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 &= 2 \\ \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 &= 3 \\ \alpha_3 + \alpha_4 + \alpha_5 &= 2 \\ \alpha_4 + \alpha_5 &= 5 \\ \alpha_5 &= -2 \end{aligned}\right\} \Rightarrow \begin{aligned} \alpha_1 &= -1 \\ \alpha_2 &= 1 \\ \alpha_3 &= -3 \\ \alpha_4 &= 7 \\ \alpha_5 &= -2 \end{aligned}$$

por lo que entonces

$$(2 + 3x + 2x^2 + 5x^3 - 2x^4)_\beta = (-1, 1, -3, 7, -2) \in \mathbf{R}^5$$

Véase un ejemplo más

EJEMPLO 2 En $M_{2\times 2}$ considere los vectores de la base canónica β_1.

$$v_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Sea

$$v = \begin{bmatrix} 2 & -3 \\ 7 & 4 \end{bmatrix}$$

Es claro entonces que

$$v = \begin{bmatrix} 2 & -3 \\ 7 & 4 \end{bmatrix} = 2v_1 - 3v_2 + 7v_3 + 4v_4$$

por lo que

$$(v)_{\beta_1} = (2, -3, 7, 4) \in \mathbf{R}^4$$

En el mismo espacio $M_{2\times 2}$, considere los vectores

$$u_1 = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 1 & 3 \\ 1 & 0 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}, \quad u_4 = \begin{bmatrix} 0 & 0 \\ -2 & -3 \end{bmatrix}$$

Se puede verificar que $\beta_2 = \{u_1, u_2, u_3, u_4\}$ es también una base de $M_{2\times 2}$.

Al tomar el mismo vector

$$v = \begin{bmatrix} 2 & -3 \\ 7 & 4 \end{bmatrix}$$

se encontrará $(v)_{\beta_2}$.

Se tiene entonces que

$$\begin{bmatrix} 2 & -3 \\ 7 & 4 \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 & 3 \\ 1 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix} + \alpha_4 \begin{bmatrix} 0 & 0 \\ -2 & -3 \end{bmatrix}$$

Al realizar las operaciones indicadas se obtiene

$$\begin{aligned} \alpha_1 + \alpha_2 \qquad\qquad &= 2 \\ 2\alpha_1 + 3\alpha_2 + \alpha_3 \qquad &= -3 \\ \alpha_2 + 2\alpha_3 - 2\alpha_4 &= 7 \\ 2\alpha_3 - 3\alpha_4 &= 4 \end{aligned}$$

Al proceder por el método de eliminación Gaussiana para resolver este sistema se tiene

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 2 \\ 2 & 3 & 1 & 0 & -3 \\ 0 & 1 & 2 & -2 & 7 \\ 0 & 0 & 2 & -3 & 4 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -25 \\ 0 & 1 & 0 & 0 & 27 \\ 0 & 0 & 1 & 0 & -34 \\ 0 & 0 & 0 & 1 & -24 \end{bmatrix}$$

de modo que

$$v = \begin{bmatrix} 2 & -3 \\ 7 & 4 \end{bmatrix} = -25 \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} + 27 \begin{bmatrix} 1 & 3 \\ 1 & 0 \end{bmatrix} - 34 \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix} - 24 \begin{bmatrix} 0 & 0 \\ -2 & -3 \end{bmatrix}$$

y entonces

$$(v)_{\beta_2} = (-25, 27, -34, -24) \in \mathbf{R}^4$$

Es natural preguntarse si existe alguna relación entre los vectores

$$(v)_{\beta_1} = (2, -3, 7, 4)$$

$$(v)_{\beta_2} = (-25, 27, -34, -24)$$

pues en realidad éstos representan *el mismo* vector $v \in M_{2\times 2}$ sólo que en diferentes bases.

El objetivo de la siguiente subsección es discutir, en general, esta pregunta.

6.1. CAMBIOS DE BASE

Sea V un espacio vectorial de dimensión n y sean

$$\beta_1 = \{v_1, v_2, \ldots, v_n\} \qquad \text{y} \qquad \beta_2 = \{u_1, u_2, \ldots, u_n\}$$

dos bases distintas de V.

Tómese un vector $v \in V$. Para este vector existen dos representaciones diferentes de acuerdo a cada una de las bases β_1 y β_2, dígase

$$v = \varphi_1 u_1 + \varphi_2 u_2 + \dots + \varphi_n u_n \tag{6.2}$$

$$v = \delta_1 v_1 + \delta_2 v_2 + \dots + \delta_n v_n \tag{6.1}$$

o sea que

$$[v]_{\beta_1} = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{bmatrix} \qquad [v]_{\beta_2} = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{bmatrix}$$

Se quiere ver qué relación existe entre $[v]\beta_1$ y $[v]\beta_2$.

Como β_1 es una base de V, cada vector $u_j, j = 1, 2, \dots, n$ (de la base β_2) puede expresarse de manera única como una combinación lineal de los vectores de β_1. Es decir, existen escalares p_{ij} tales que

$$u_j = p_{1j}v_1 + p_{2j}v_2 + \dots + p_{nj}v_n = \sum_{i=1}^{n} p_{ij}v_i \qquad j = 1, 2, \dots, n \tag{6.3}$$

o sea que

$$[u_j]_{\beta_1} = \begin{bmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{nj} \end{bmatrix} \qquad j = 1, 2, \dots, n$$

Sustitúyase (6.3) en (6.2) para obtener

$$v = \varphi_1 u_1 + \varphi_2 u_2 + \dots + \varphi_n u_n = \sum_{j=1}^{n} \varphi_j u_j = \sum_{j=1}^{n} \varphi_j \sum_{i=1}^{n} p_{ij} v_i =$$

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{ij}\,\varphi_j \right) v_i = \left(\sum_{j=1}^{n} p_{1j}\,\varphi_j \right) v_1 + \left(\sum_{j=1}^{n} p_{2j}\,\varphi_j \right) v_2 + \dots + \left(\sum_{j=1}^{n} p_{nj}\,\varphi_j \right) v_n \tag{6.4}$$

Al comparar (6.4) con (6.1) y en vista de la unicidad de la representación de v como combinación lineal de los vectores de la base β_1, se obtiene

$$\delta_1 = \sum_{j=1}^{n} p_{1j}\,\varphi_j, \quad \delta_2 = \sum_{j=1}^{n} p_{2j}\,\varphi_j, \quad \dots, \quad \delta_n = \sum_{j=1}^{n} p_{nj}\,\varphi_j$$

o sea

$$\delta_i = \sum_{j=1}^{n} p_{ij}\varphi_j \tag{6.5}$$

Considérese la matriz $P = (p_{ij})_{i,j=1,2,\dots,n}$, en cuya j-ésima columna se encuentran los elementos de la matriz $[u_j]_{\beta_1}$.

$$P = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ & & \dots & \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix}$$

o esquemáticamente

$$P = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ [u_1]_{\beta_1} & [u_2]_{\beta_1} & \dots & [u_n]_{\beta_1} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

Obsérvese que

$$P[v]_{\beta_2} = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ & & \dots & \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} p_{1j}\varphi_j \\ \sum_{j=1}^{n} p_{2j}\varphi_j \\ \vdots \\ \sum_{j=1}^{n} p_{nj}\varphi_j \end{bmatrix} = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{bmatrix} = [v]_{\beta_1}$$

Es decir, la matriz de coordenadas de v respecto de la base β_1 se obtiene multiplicando la matriz P por la matriz de coordenadas de v respecto de la base β_2.

Si se intercambian los papeles de β_1 y β_2 se obtiene la siguiente situación "dual"

$$Q[v]_{\beta_1} = [v]_{\beta_2} \tag{6.6}$$

en donde ahora la matriz

$$Q = \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ & & \dots & \\ q_{n1} & q_{n2} & \dots & q_{nn} \end{bmatrix}$$

es la que tiene en su i-ésima columna los elementos de la matriz de coordenadas del i-ésimo vector v_i de la base β_1 respecto de la base β_2, es decir,

$$Q = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ [v_1]_{\beta_2} & [v_2]_{\beta_2} & \dots & [v_n]_{\beta_2} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \quad [v_i]_{\beta_2} = \begin{bmatrix} q_{1i} \\ q_{2i} \\ \vdots \\ q_{ni} \end{bmatrix} \tag{6.7}$$

Se afirma que $QP = I$. En efecto, según (6.3) se tiene

$$u_j = \sum_{i=1}^{n} p_{ij} v_i$$

pero según (6.7)

$$v_i = \sum_{k=1}^{n} q_{ki} u_k$$

de modo que

$$u_j = \sum_{i=1}^{n} p_{ij} v_i = \sum_{i=1}^{n} p_{ij} \sum_{k=1}^{n} q_{ki} u_k = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} q_{ki} p_{ij} \right) u_k \qquad (6.8)$$

Pero por otra parte, es claro que el vector u_j se expresa como combinación lineal de los vectores de la base β_2 como

$$u_j = \sum_{k=1}^{n} \gamma_k u_k \qquad \text{en donde} \qquad \gamma_k = \begin{cases} 1 \text{ si } k = j \\ 0 \text{ si } k \neq j \end{cases} \qquad (6.9)$$

de modo que, al comparar (6.8) y (6.9) se obtiene

$$\sum_{i=1}^{n} q_{ki} p_{ij} = \begin{cases} 1 \text{ si } k = j \\ 0 \text{ si } k \neq j \end{cases} \qquad (6.10)$$

Pero $\sum_{i=1}^{n} q_{ki} p_{ij}$ es el elemento que aparece en la k-ésima línea y en la j-ésima columna de la matriz producto QP.

Entonces (6.10) dice que $QP = I$.

Se concluye entonces que P es inversible y que $P^{-1} = Q$.

DEFINICIÓN 6.2 A la matriz P de orden n tal que

$$P[v]_{\beta_2} = [v]_{\beta_1}$$

se le llama *matriz de cambio de base de* β_2 *a* β_1.

Se ha visto entonces que P es una matriz inversible y que P^{-1} es la matriz de cambio de base de β_1 a β_2, es decir

$$[v]_{\beta_2} = P^{-1}[v]_{\beta_1}$$

En el siguiente teorema se resume (sin palabras) la discusión anterior.

TEOREMA 6.1 Sea V un espacio vectorial de dimensión n.
Sean

$$\beta_1 = \{v_1, v_2, \ldots, v_n\}$$
$$\beta_2 = \{u_1, u_2, \ldots, u_n\}$$

dos bases de V. Sean P y Q las matrices

$$P = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ [u_1]_{\beta_1} & [u_2]_{\beta_1} & \cdots & [u_n]_{\beta_1} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \qquad Q = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ [v_1]_{\beta_2} & [v_2]_{\beta_2} & \cdots & [v_n]_{\beta_2} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

Entonces P (es inversible y) es la matriz de cambio de base de β_2 a β_1, y $Q = P^{-1}$ es la matriz de cambio de base de β_1 a β_2.

Esquemáticamente, se puede imaginar a P como siendo una "máquina que convierte vectores de coordenadas respecto de la base β_2 a vectores de coordenadas respecto de la base β_1".

$$[v]_{\beta_2} \rightarrow \boxed{P} \rightarrow [v]_{\beta_1} = P[v]_{\beta_2}$$

CAMBIO DE
BASE $\beta_2 \rightarrow \beta_1$

Similarmente para P^{-1} (la misma "máquina" con las flechas en sentido inverso).

EJEMPLO 3 Retome el ejemplo del espacio $M_{2\times 2}$ en el que

$$\beta_1 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

$$\beta_2 = \left\{ \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ -2 & -3 \end{bmatrix} \right\}$$

Es muy fácil hallar la matriz de cambio de base de β_1 a β_2, pues

$$\left(\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \right)_{\beta_1} = (1, 2, 0, 0) \qquad \left(\begin{bmatrix} 1 & 3 \\ 1 & 0 \end{bmatrix} \right)_{\beta_1} = (1, 3, 1, 0)$$

$$\left(\begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix} \right)_{\beta_1} = (0, 1, 2, 2) \qquad \left(\begin{bmatrix} 0 & 0 \\ -2 & -3 \end{bmatrix} \right)_{\beta_1} = (0, 0, -2, -3)$$

de modo que

$$P = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 2 & -3 \end{bmatrix}$$

Se puede hallar Q la matriz de cambio de base de β_1 a β_2, calculando P^{-1}. Ésta es

$$Q = P^{-1} = \begin{bmatrix} -3 & 2 & -3 & 2 \\ 4 & -2 & 3 & -2 \\ -6 & 3 & -3 & 2 \\ -4 & 2 & -2 & 1 \end{bmatrix}$$

(o bien, expresando directamente cada vector de la base β_1 como combinación lineal de los vectores de la base β_2 y escribiendo las coordenadas en las columnas correspondientes de P).

Para el vector

$$v = \begin{bmatrix} 2 & -3 \\ 7 & 4 \end{bmatrix}$$

se tiene

$$[v]_{\beta_1} = \begin{bmatrix} 2 \\ -3 \\ 7 \\ 4 \end{bmatrix}$$

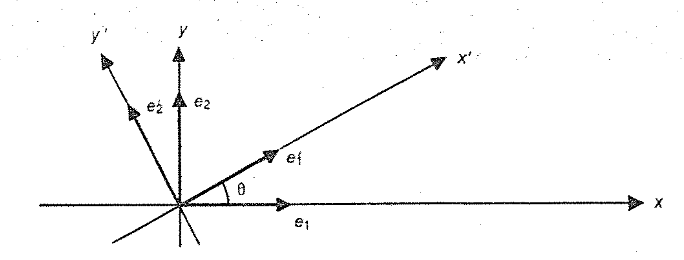
de modo que

$$[v]_{\beta_2} = Q[v]_{\beta_1} = \begin{bmatrix} -3 & 2 & -3 & 2 \\ 4 & -2 & 3 & -2 \\ -6 & 3 & -3 & 2 \\ -4 & 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 7 \\ 4 \end{bmatrix} = \begin{bmatrix} -25 \\ 27 \\ -34 \\ -24 \end{bmatrix}$$

o sea que $[v]_{\beta_2} = (-25, 27, -34, -24)$, resultado que ya se había obtenido directamente.

Para terminar, se desearía considerar un ejemplo geométrico muy interesante, con el cual es muy probable que el lector tenga familiaridad. Se trata de estudiar la "rotación de ejes" en el plano xy.

Estúdiese la situación geométrica descrita en la siguiente figura en la que se muestra el plano xy y un nuevo sistema de coordenadas $x'y'$ que se obtiene girando los ejes xy. Un ángulo θ.



Sea $\beta_1 = \{e_1, e_2\}$ la **base canónica** de $\mathbb{R}^2$.

Sean $e_1' = (\cos\theta, \operatorname{sen}\theta)$ y $e_2' = (-\operatorname{sen}\theta, \cos\theta)$. Es fácil verificar que para cualquier ángulo θ, los vectores e_1 y e_2 constituyen también una base de $\mathbb{R}^2$. Escríbase $\beta_2 = \{e_1', e_2'\}$. (Obsérvese que ésta sería la base canónica en el plano $x'y'$.) Dado un vector $v = (x, y) \in \mathbb{R}^2$, se quiere ver cuál es la relación que existe entre las coordenadas de v en la base β_1 [que son precisamente (x, y)] y sus coordenadas en la base β_2 (en el sistema xy "girado").

Como

$$(e_1')_{\beta_1} = (\cos\theta, \operatorname{sen}\theta)$$

$$(e_2')_{\beta_1} = (-\operatorname{sen}\theta, \cos\theta)$$

la matriz de cambio de base de β_2 a β_1, es entonces

$$P = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

cuya inversa es

$$P^{-1} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

que es entonces la matriz de cambio de base de β_1 a β_2. Sean (x', y') las coordenadas de v con respecto a la base β_2.

Entonces

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = P^{-1}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta + y\operatorname{sen}\theta \\ -x\operatorname{sen}\theta + y\cos\theta \end{bmatrix}$$

es decir

$$x' = x\cos\theta + y\operatorname{sen}\theta$$
$$y' = -x\operatorname{sen}\theta + y\cos\theta$$

que son las conocidas "fórmulas de cambio de coordenadas por rotación de ejes", que relacionan las coordenadas de un vector (x, y) en el sistema xy con las coordenadas (x', y') de ese mismo vector en el sistema girado $x'y'$.

6.2 ISOMORFISMOS DE ESPACIOS VECTORIALES

En esta subsección se pretende ir un poco más lejos respecto de las consecuencias teóricas que trae consigo la identificación entre los vectores de un espacio vectorial V de dimensión n y sus vectores de coordenadas en el espacio $\mathbf{R}^n$.

En la siguiente definición se establece uno de los conceptos más importantes del álgebra lineal: el concepto de **isomorfismo.**

DEFINICIÓN 6.3 Sean V y U dos espacios vectoriales. Se dice que la función $f: V \to U$ es un *isomorfismo* de V a U si f es biyectiva (inyectiva y sobreyectiva) y además cumple con las dos condiciones siguientes:

1) $f(v_1 + v_2) = f(v_1) + f(v_2) \quad \forall v_1, v_2 \in V.$
2) $f(cv) = cf(v), \quad \forall c \in \mathbf{R}, \ \forall v \in V.$

NOTA: A una función $f: V \to U$ entre los espacios vectoriales V y U que satisface las condiciones 1) y 2) de la definición anterior se le llama "función (o transformación) lineal". Éstas juegan un papel central dentro del estudio del álgebra lineal. Todo el capítulo 4 se dedicará a hacer un estudio de estas importantes funciones.

En el siguiente teorema se establecen dos propiedades muy importantes de los isomorfismos que se usarán más adelante.

TEOREMA 6.2 Sea $f: V \to U$ un **isomorfismo** del espacio vectorial V al espacio vectorial U. Entonces

a) $f^{-1}: U \to V$ es un isomorfismo de U a V.
b) si $g: U \to W$ es un isomorfismo de U a W, la composición $g \circ f: V \to W$ es también un isomorfismo (de U a W).

DEMOSTRACIÓN a) Siendo $f: V \to U$ una biyección, su inversa $f^{-1}: U \to V$ también es una función biyectiva. Se verá que f^{-1} satisface las condiciones 1) y 2) de la definición 6.3. Se tienen que verificar que

$$f^{-1}(u_1 + u_2) = f^{-1}(u_1) + f^{-1}(u_2) \qquad \forall u_1, u_2 \in U$$

y

$$f^{-1}(cu) = cf^{-1}(u) \qquad \forall c \in \mathbf{R} \qquad \forall u \in U$$

Sean $u_1, u_2 \in U$. Sean $v_1, v_2 \in V$ tales que $f(v_1) = u_1$ y $f(v_2) = u_2$ (esto es, $u_1 = f^{-1}(v_1)$ y $u_2 = f^{-1}(v_2)$). Entonces

$$u_1 + u_2 = f(v_1) + f(v_2) = f(v_1 + v_2) \qquad \text{(pues } f \text{ es un isomorfismo)}$$

de donde

$$f^{-1}(u_1 + u_2) = v_1 + v_2 = f^{-1}(u_1) + f^{-1}(u_2)$$

Similarmente, si $c \in \mathbf{R}$ y $u \in U$, tome $v \in V$ tal que $f(v) = u$.
Entonces

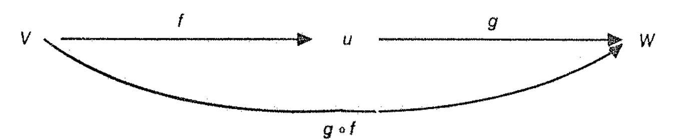
$$cu = cf(v) = f(cv) \qquad \text{(pues } f \text{ es un isomorfismo)}$$

o sea

$$f^{-1}(cu) = cv = cf^{-1}(u)$$

lo que prueba entonces a).

b) Se tiene la siguiente situación:



Siendo f y g biyecciones, la composición $g \circ f$ también es una biyección. Se verá entonces si $g \circ f$ satisface las condiciones l) y 2) de la definición 6.3.

Sean $v_1, v_2 \in V$. Entonces

$$(g \circ f)(v_1 + v_2) = g(f(v_1 + v_2)) = g(f(v_1) + f(v_2)) = g(f(v_1)) + g(v(v_2)) = (g \circ f)(v_1) + (g \circ f)(v_2)$$

Similarmente, si $v \in V$ y $c \in R$

$$(g \circ f)(cv) = g(f(cv)) = g(cf(v)) = cg(f(v)) = c(g \circ f)(v)$$

lo que prueba (b).

Q.E.D.

Considérese ahora a la familia de todos los espacios vectoriales. Establézcase en esta familia la siguiente relación:

$$V \sim U \Leftrightarrow \text{existe un isomorfismo } f: V \to U.$$

TEOREMA 6.3 La relación anterior es una relación de equivalencia.

DEMOSTRACIÓN 1) La relación ~ es reflexiva. Es decir $V \sim V$ para todo espacio vectorial V. Esto es claro, pues la función identidad $Id: V \to V$, $Id(v) = v \ \forall v \in V$, es un isomorfismo de V a V.

2) La relación ~ es simétrica. Es decir, si $V \sim U$ entonces $U \sim V$. En efecto, si $V \sim U$, existe un isomorfismo f: $V \to U$. Por el teorema 6.2, f^{-1}: $U \to V$ es un isomorfismo de U a V. Entonces $U \sim V$.

3) La relación ~ es transitiva. Es decir, si $V \sim U$ y $U \sim W$, entonces $V \sim W$. En efecto, si $V \sim U$ existe un isomorfismo f: $V \to U$. Ahora, como $U \sim W$, existe un isomorfismo g: $U \to W$. El teorema 6.2 dice que $g \circ f$: $V \to W$ es un isomorfismo de V a W. Entonces $V \sim W$.

Q.E.D.

En vista de este teorema (de la propiedad simétrica de la relación ~, más concretamente) se podrá referir a un isomorfismo f: $V \to U$ de V a U, simplemente como un isomorfismo "entre V y U" sin temor a confusión.

DEFINICIÓN 6.4 Se dice que los espacios vectoriales U y V son *isomorfos*, lo cual se escribe $U \cong V$, si existe un isomorfismo entre ellos.

En vista del teorema 6.3, se tiene entonces que la familia de todos los espacios vectoriales queda dividida en clases de equivalencia. Todos los espacios que pertenecen a una misma clase son isomorfos. Esto permite *identificar* dos espacios isomorfos como "siendo el mismo"* (identificarlos con su clase de equivalencia). El siguiente teorema dice que muchos de los espacios vectoriales que se han estudiado en este capítulo pertenecen a una misma clase de equivalencia.

TEOREMA 6.4 Sea V un espacio vectorial de dimensión n. Entonces $V \cong \mathbf{R}^n$.

DEMOSTRACIÓN Fíjese una base β del espacio vectorial V. Dígase que $\beta = \{v_1, v_2, \ldots v_n\}$.

Defínase la función f: $V \to \mathbf{R}^n$ de la siguiente manera: para $v \in V$ escriba $f(v) = (v)_\beta$. (El vector de coordenadas de v respecto de la base β.)

Se afirma que f es un isomorfismo entre V y R^n. El hecho de que f es inyectiva, está establecido en el teorema 5.1 (unicidad de la representación de cada vector $v \in V$ como combinación lineal de los elementos de la base β). La función f también es sobreyectiva: dado el vector $(c_1, c_1, \ldots, c_n) \in \mathbf{R}^n$ defina $v \in V$ escribiendo

$$v = c_1v_1 + c_2v_2 + \ldots + c_nv_n$$

Es claro que $f(v) = (c_1, c_2, \ldots, c_n)$.

Se verá ahora que f satisface las condiciones 1) y 2) de la definición 6.3. Sean $v, u, \in V$, dígase que

$$v = \tau_1v_1 + \tau_2v_2 + \ldots + \tau_nv_n$$

$$u = \mu_1v_1 + \mu_2v_2 + \ldots + \mu_nv_n$$

*La palabra **isomorfismo** tiene por raíces "isos" = igual, "morfos" = forma.

Entonces

$$v + u = (\tau_1 + \mu_1)v_1 + (\tau_2 + \mu_2)v_2 + \ldots + (\tau_n + \mu_n)v_n$$

Entonces

$$f(v + u) = (\tau_1 + \mu_1, \tau_2 + \mu_2, \ldots, \tau_n + \mu_n)$$
$$= (\tau_1, \tau_2, \ldots, \tau_n) + (\mu_1, \mu_2, \ldots, \mu_n)$$
$$= f(v) + f(u)$$

Si $c \in \mathbf{R}$ entonces

$$cv = c\tau_1v_1 + c\tau_2v_2 + \ldots + c\tau_nv_n$$

y por tanto

$$f(cv) = (c\tau_1, c\tau_2, \ldots, c\tau_n) = c(\tau_1, \tau_2, \ldots, \tau_n) = cf(v)$$

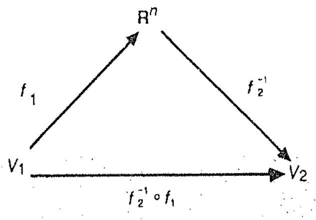
Q.E.D.

COROLARIO 1 $P_n \cong \mathbf{R}^{n+1}$, $M_{m \times n} \cong \mathbf{R}^{mn}$.

COROLARIO 2 Si V_1 y V_2 son dos espacios vectoriales de la misma dimensión entonces $V_1 \cong V_2$.

DEMOSTRACIÓN Sea $n = dim\ V_1 = dim\ V_2$. Según el teorema anterior existen isomorfismos $f_1 = V_1 \to \mathbf{R}^n$ y $f_2 = V_2 \to \mathbf{R}^n$.

Considérese la función $f_2^{-1} \circ f_1$: $V_1 \to V_2$. Esquemáticamente se tiene la siguiente situación:



Según el teorema 6.2, $f_2^{-1} \circ f_1$ es un isomorfismo. Esto prueba entonces que $V_1 \cong V_2$.*

Q.E.D.

*Una demostración más directa de este corolario (usando el teorema 6.3) es: si $n = dim\ V_1 = dim\ V_2$, se tiene $V_1 \cong \mathbf{R}^n \cong V_2$, y por tanto $V_1 \cong V_2$. Se quiso aquí dar explícitamente el isomorfismo entre V_1 y V_2 pues nos referiremos a él en la discusión que sigue a este corolario.

Para establecer el isomorfismo $f_2^{-1} \circ f_1$: $V_1 \to V_2$ en el corolario anterior es necesario fijar bases de V_1 y V_2, dígase β_1 de V_1 y β_2 de V_2. Este isomorfismo sería entonces (véase demostración del teorema 6.4)

$$(f_2^{-1} \circ f_1)(v) = c_1v_1 + c_2v_2 + \ldots + c_nv_n$$

en donde $\beta_2 = \{v_1, v_2, \ldots, v_n\}$ y $(v)_{\beta_1} = (c_1, c_2, \ldots, c_n)$

En particular, para identificar a los vectores de P_n con los vectores R^{n+1}, se podría hacer por medio de una base concreta β de P_n (y la base canónica de $\mathbf{R}^{n+1}$), asociando a cada vector $v \in P_n$ su vector de coordenadas $(v)_\beta \in \mathbf{R}^{n+1}$. Al tomar β como la base canónica de P_n, esta identificación se ve como

$$a_0 + a_1x + \ldots + a_nx^n \in P_n \leftrightarrow (a_0, a_1, \ldots, a_n) \in \mathbf{R}^{n+1}$$

Similarmente con $M_{m\times n}$ se tiene la siguiente identificación vía la base canónica de $M_{m\times n}$.

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & \ldots & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \in M_{m\times n}$$

$$\updownarrow$$

$$(a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, a_{22}, \ldots, a_{2n}, \ldots, a_{m1}, a_{m2}, \ldots, a_{mn}) \in \mathbf{R}^{mn}$$

El ocasiones se suele referir a esta identificación como el "isomorfismo natural" entre $P_n(M_{m\times n})$ y $\mathbf{R}^{n+1}$ ($\mathbf{R}^{mn}$ respectivamente).

Ya en un par de ocasiones anteriores se ha hecho la identificación de una matriz $n \times 1$ con un vector de $\mathbf{R}^n$ (por ejemplo cuando se trata sobre el espacio de soluciones de un sistema homogéneo de ecuaciones lineales, en la sección 3).

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in M_{n\times 1} \quad \leftrightarrow \quad (a_1, a_2, \ldots, a_n) \in \mathbf{R}^n$$

Esta identificación no es peligrosa en modo alguno, pues ya se vio que $M_{n\times 1} \simeq \mathbf{R}^n$.

Una de las características importantes que tienen los espacios vectoriales V y W que son isomorfos, es que algunas propiedades importantes —bajo la perspectiva del álgebra lineal— que posee un conjunto de vectores en uno de los espacios, se conservan a través del isomorfismo hacia el otro espacio. Más concretamente, si f: $V \to W$ es un isomorfismo se tiene

1) S es un conjunto que genera a V si, y sólo si $f(S)$ es un conjunto que genera a W.

2) S es un conjunto linealmente independiente en V si, y sólo si $f(S)$ es un conjunto linealmente independiente en W.

y por tanto,

3) β es una base de V si, y sólo si $f(\beta)$ es una base de W.

La demostración de estos hechos no es difícil. Sin embargo, se dejará para el próximo capítulo (sección 6) en donde se retomará la discusión que sobre isomorfismos se hizo en esta subsección. Por lo pronto, en la próxima sección se explotarán estas propiedades para establecer un método sencillo que permita encontrar bases de subespacios generados por un conjunto de vectores en un espacio vectorial de dimensión n.

Para terminar la discusión que se ha presentado en esta subsección sobre isomorfismos, se desea hacer algunos comentarios importantes sobre la perspectiva que se abre a la luz de los resultados aquí obtenidos. Se tiene entonces que todos los espacios vectoriales de dimensión n son isomorfos. Si dos espacios vectoriales V_1 y V_2 son isomorfos, dígase que f: $V_1 \to V_2$ es un isomorfismo, la estructura de espacio vectorial de V_1 es "copiada fielmente" por f hacia V_2, pues por una parte f establece una correspondencia biyectiva entre los elementos de V_1 y V_2 (son identificados uno a uno) y por otra parte f traslada sumas y productos por escalares —las operaciones de espacio vectorial— de los elementos de V_1 a sumas y productos por escalares de los correspondientes elementos de V_2. Así pues, desde el punto de vista del álgebra lineal, V_1 y V_2 son espacios vectoriales indistinguibles. Se llega entonces a la conclusión de que todos los espacios vectoriales de dimensión n son, bajo la óptica del álgebra lineal, indistinguibles respecto del espacio $\mathbf{R}^n$. ¿Por qué no estudiar entonces solamente el espacio $\mathbf{R}^n$?

Existen muchas razones por las cuales esto no es conveniente.

Sin pretender entrar en una discusión detallada sobre este respecto, véase sólo un ejemplo. Los espacios $M_{n\times n}$ y $\mathbf{R}^{n^2}$ son isomorfos, pues ambos tienen dimensión n^2. Supóngase que se quiere estudiar el espacio $M_{n\times n}$ vía el espacio $\mathbf{R}^{n^2}$ por medio del isomorfismo natural entre estos espacios. (Primera objeción: se han comprometido con bases concretas de $M_{n\times n}$ y $\mathbf{R}^{n^2}$, las bases canónicas; ¿por qué no otras bases?) Existen algunas propiedades de los vectores en el espacio $M_{n\times n}$ que tienen un carácter exclusivo para este espacio. Estas propiedades hacen del espacio $M_{n\times n}$, a pesar de su isomorfismo con $\mathbf{R}^{n^2}$, un espacio con características propias y exclusivas. Piense, por ejemplo, en la multiplicación de matrices, la inversibilidad, los determinantes de estas matrices, etc. Debe resultar claro el panorama tan tenebroso que se presentaría si se pretendiera establecer estas propiedades entre las matrices vía sus imágenes en $\mathbf{R}^{n^2}$.

La identificación proporcionada por los isomorfismos entre espacios vectoriales de la misma dimensión resulta ser importante en ciertos niveles del álgebra lineal que están fuera de los alcances de este libro. Así entonces, para uno seguirán siendo "distintos" los vectores de $M_{m\times n}$, P_{mn-1} y $\mathbf{R}^{m\times n}$, aunque estos espacios sean isomorfos. Solamente se seguirán identificando las matrices de orden $n \times 1$ con su correspondiente vector de coordenadas (respecto de la base canónica) en $\mathbf{R}^n$. Es claro que esto se hace sólo por una cuestión de comodidad en la escritura.

EJERCICIOS (SECCIÓN 6, CAPÍTULO 3)

1. a) Sea β la base canónica de $\mathbf{R}^n$ y sea x un vector cualquiera de este espacio, encuentre $(x)_\beta$.
   b) Sea β la base canónica de $M_{m\times n}$ y sea A una matriz cualquiera de este espacio, encuentre $(A)_\beta$.
   c) Sea β la base canónica de P_n y sea p un polinomio cualquiera de este espacio, encuentre $(p)_\beta$.
2. Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base cualquiera del espacio vectorial V,
   a) ¿cuál es el vector de coordenadas del vector cero de V respecto de la base β?
   b) ¿cuál es el vector de coordenadas del vector v_i $i = 1, 2, \ldots, n$ respecto de la base β?
3. Considere el conjunto de vectores en $\mathbf{R}^3$

$$\beta = \{(3, 1, 2), (-1, 0, 2), (4, 3, 5)\}$$

   a) Demuestre que β es una base de $\mathbf{R}^3$.
   b) Encuentre $((1, 1, 1))_\beta$ y $((3, 4, 6))_\beta$.
   c) Para un vector cualquiera $(x, y, z) \in \mathbf{R}^3$, encuentre $((x, y, z))_\beta$.
4. Considere el conjunto de vectores en P_3

$$\beta = \{1 + x + x^2 + x^3, 2 + x^2 - x^3, 3x + x^3, 2 + 3x - x^2\}$$

   a) Compruebe que β es una base de P_3.
   b) Encuentre $(1)_\beta$, $(x)_\beta$, $(x^2)_\beta$ y $(x^3)_\beta$.
5. Sea β una base del espacio vectorial V. Sean v_1 y v_2 dos vectores de V y c un escalar, demuestre que

$$(v_1 + v_2)_\beta = (v_1)_\beta + (v_2)_\beta$$
$$(cv_1)_\beta = c(v_1)_\beta$$

6. Use el ejercicio anterior para hallar

$$(a_0 + a_1x + a_2x^2 + a_3x^3)_\beta$$

   en donde β es la base de P_3 del ejercicio 4.
7. Sea $\beta = \{v_1, v_2\}$ una base de $\mathbf{R}^2$, suponga que

$$((1, 1))_\beta = (2, 3) \qquad ((-1, 1))_\beta = (3, 5)$$

   Determine los vectores de la base β.
8. Sea $\beta = \{v_1, v_2, v_3\}$ una base de P_2, suponga que

$$(1 + x)_\beta = (2, 1, 0) \qquad (2 - x + x^2)_\beta = (1, 1, 1) \qquad (x - x^2)_\beta = (-1, 3, 2)$$

   Determine los vectores de la base β.
9. Encuentre el vector de coordenadas del vector

$$p(x) = a_0 + a_1x + \ldots + a_nx^n$$

   del espacio P_n en términos de la base

$$\beta = \{1, x - 1, (x - 1)^2, \ldots, (x - 1)^n\}$$

   (*Sugerencia*: el resultado se expresa fácilmente en términos de las derivadas del polinomio $p(x)$ evaluadas en $x = 1$.)
10. En $\mathbf{R}^3$ considere las dos bases

$$\beta_1 = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$$
$$\beta_2 = \{(2, 1, 2), (1, 0, 3), (-1, 4, -2)\}$$

   a) Encuentre $((1, 1, 3))_{\beta_1}$.
   b) Halle $((1, 1, 3))_{\beta_2}$.
   c) Encuentre la matriz de cambio de base de β_1 a β_2.
   d) Halle la matriz de cambio de base de β_2 a β_1.
   e) Reobtenga los resultados de los incisos a) y b) usando las matrices obtenidas en los incisos c) y d).
11. En $\mathbf{R}^4$ considere las dos bases

$$\beta_1 = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$$
$$\beta_2 = \{(1, 0, 0, 0), (2, 2, 0, 0), (3, 3, 3, 0), (4, 4, 4, 4)\}$$

   a) Encuentre la matriz de cambio de base de β_2 a β_1.
   b) Halle la matriz de cambio de base de β_1 a β_2.
   c) Use la matriz obtenida en el inciso anterior para hallar las coordenadas de un vector cualquiera (a, b, c, d) en términos de la base β_2.
12. En $\mathbf{R}^5$ considere las dos bases

$$\beta_1 = \{(1, 0, 0, 0, 0), (0, 1, 0, 0, 0), (0, 0, 1, 0, 0), (0, 0, 0, 1, 0), (0, 0, 0, 0, 1)\}$$
$$\beta_2 = \{(0, 0, 0, 0, 1), (0, 0, 0, 1, 0), (0, 0, 1, 0, 0), (0, 1, 0, 0, 0), (1, 0, 0, 0, 0)\}$$

   a) Encuentre la matriz de cambio de base de β_2 a β_1.
   b) Halle la matriz de cambio de base de β_1 a β_2.
   c) Encuentre las coordenadas de un vector cualquiera $x = (x_1, x_2, x_3, x_4, x_5)$ en términos de la base β_2.
13. Sea V un espacio vectorial de dimensión 4, considere las dos bases de V

$$\beta_1 = \{v_1, v_2, v_3, v_4\} \qquad \beta_2 = \{u_1, u_2, u_3, u_4\}$$

   en donde

$$u_1 = 3v_1 + 5v_3 + 2v_4$$
$$u_2 = 3v_1 + 6v_2 + 4v_3 + 3v_4$$
$$u_3 = -4v_1 + v_2 + 2v_3 + 3v_4$$
$$u_4 = -3v_1 + v_2 + v_3 + 2v_4$$

   Halle la matriz de cambio de base de β_2 a β_1.
14. En el espacio vectorial $F(\mathbf{R})$, considere el subespacio V generado por sen x y cos x.
   a) Demuestre que $\beta_1 = \{\text{sen } x, \cos x\}$ es una base de V.
   b) Compruebe que $\beta_2 = \{\text{sen}(x + \pi/4), \cos(x + \pi/4)\}$ es otra base de V.
   c) Encuentre la matriz de cambio de base de β_2 a β_1.

d) Halle la matriz de cambio de base de β_1 a β_2.

15. En el espacio vectorial $F(\mathbf{R})$ considere el subespacio V generado por $\operatorname{sen}^2 x$ y $\cos^2 x$.
   a) Demuestre que $\beta_1 = \{\operatorname{sen}^2 x, \cos^2 x\}$ es una base de V.
   b) Compruebe que $\beta_2 = \{1, \cos 2x\}$ es otra base de V.
   c) Encuentre la matriz de cambio de base de β_2 a β_1.
   d) Halle la matriz de cambio de base de β_1 a β_2.

① 16. Sea V un espacio vectorial de dimensión n. Considere las dos bases de V dadas por

$$\beta_1 = \{v_1, v_2, \ldots, v_n\} \qquad \beta_2 = \{u_1, u_2, \ldots, u_n\}$$

Sea P la matriz de cambio de base de β_1 a β_2.
   a) ¿Cómo se afecta P si en la base β_1 se intercambia la posición de dos de sus vectores?
   b) ¿Cómo se afecta P si en la base β_2 se intercambia la posición de dos de sus vectores?
   c) Sean β'_1 y β'_2 las bases de V dadas por

$$\beta'_1 = \{v_n, v_{n-1}, \ldots, v_1\} \qquad \beta'_2 = \{u_n, u_{n-1}, \ldots, u_1\}$$

Sea Q la matriz de cambio de base de β'_1 a β'_2, demuestre que $Q = Pc$, en donde Pc es la matriz definida en el ejercicio 17 de la sección 2 del capítulo 2.

17. Considere el espacio vectorial P_2 con la base β dada por

$$\beta = \{1 + x - x^2, 2 + 3x + x^2, -1 - 5x + x^2\}$$

Establezca un isomorfismo $f: P_2 \to \mathbf{R}^3$ por medio de la base β.

① 18. En el espacio vectorial $M_{2\times 2}$ considere la base β dada por

$$\beta = \left\{\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}, \begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}, \begin{bmatrix}1 & -1\\ 1 & -1\end{bmatrix}, \begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\right\}$$

   a) Establezca un isomorfismo $f: M_{2\times 2} \to \mathbf{R}^4$ por medio de la base β. ¿Cuál será el isomorfismo inverso $f^{-1}: \mathbf{R}^4 \to M_{2\times 2}$?
   b) Demuestre que el conjunto

$$S = \left\{\begin{bmatrix}2 & -1\\ 3 & 0\end{bmatrix}, \begin{bmatrix}3 & 2\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 1 & -1\end{bmatrix}\right\}$$

es un conjunto linealmente independiente.
   c) Demuestre que el conjunto $f(S)$ en $\mathbf{R}^4$ dado por

$$f(S) = \left\{f\left(\begin{bmatrix}2 & -1\\ 3 & 0\end{bmatrix}\right), f\left(\begin{bmatrix}3 & 2\\ 0 & 0\end{bmatrix}\right), f\left(\begin{bmatrix}0 & 0\\ 1 & -1\end{bmatrix}\right)\right\}$$

es un conjunto linealmente independiente (y entonces el isomorfismo f transformó el conjunto linealmente independiente S en el conjunto linealmente independiente $f(S)$).

19. Establezca un isomorfismo $f: P_3 \to M_{2\times 2}$ (considere las bases canónicas en ambos espacios).

7. EL ESPACIO LÍNEA DE UNA MATRIZ

Sea la matriz A de orden $m \times n$

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ & & \ldots & \\ a_{m1} & a_{m2} & \ldots & a_{mn}\end{bmatrix}$$

Considérense los m vectores $L_i^A \in \mathbf{R}^n$ definidos por

$$L_i^A = (a_{i1}, a_{i2}, \ldots, a_{in}) \qquad i = 1, 2, \ldots, m$$

Las coordenadas del vector L_i^A coinciden con los correspondientes elementos de la i-ésima línea de la matriz A. A estos vectores se les llama (por razones obvias) *vectores línea de la matriz A*.

DEFINICIÓN 7.1 Con la notación anterior, se llama *espacio línea* de la matriz A al subespacio de $\mathbf{R}^n$ generado por los vectores línea de A.

Se denotará por EL^A al espacio línea de A, el cual es entonces

$$EL^A = \mathcal{L}(L_1^A, L_2^A, \ldots, L_m^A) \subseteq \mathbf{R}^n$$

En esta sección se propone ver cómo hallar una base para el espacio EL^A (y por lo tanto, su dimensión) y aplicar estas ideas, junto con las desarrolladas en la sección anterior, para establecer un método sencillo que permita describir los subespacios generados por un conjunto de vectores en un espacio de dimensión finita.

El primer paso que se dará en esta dirección es relacionar el espacio línea de una matriz A con el espacio línea de la forma escalonada reducida de A. Se verá en el teorema 7.1 que estos espacios son iguales.

Primero se hará una observación de carácter técnico acerca de una manera conveniente (para los propósitos en esta sección) de ver un producto de matrices.

Sea A una matriz de orden $m \times p$ con elementos a_{ij} y B una matriz de orden $p \times n$ con elementos b_{ij}. Fórmese su producto $C = AB$. C es entonces una matriz de orden $m \times n$, con elementos dígase c_{ij}.

Se sabe que

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

Entonces

$$L_i^C = (c_{i1}, c_{i2}, \ldots, c_{in}) = \left(\sum_{k=1}^{p} a_{ik} b_{k1}, \sum_{k=1}^{p} a_{ik} b_{k2}, \ldots, \sum_{k=1}^{p} a_{ik} b_{kn}\right)$$

$$= \sum_{k=1}^{p} a_{ik}(b_{k1}, b_{k2}, \ldots, b_{kn}) = \sum_{k=1}^{p} a_{ik}L_k^B$$

$$L_i^C = a_{i1}L_1^B + a_{i2}L_2^B + \ldots + a_{ip}L_p^B$$

Esta expresión nos dice entonces que el i-ésimo vector línea de la matriz producto $C = AB$ es una combinación lineal de los vectores línea de la matriz B, en la cual los coeficientes de estos vectores son los correspondientes elementos de la i-ésima línea de A.

EJEMPLO 1 Por ejemplo:

$$\underset{A}{\begin{bmatrix} 2 & -3 & 1 \\ 4 & -1 & 2 \end{bmatrix}} \underset{B}{\begin{bmatrix} 2 & 1 & 3 \\ -1 & 0 & 1 \\ 3 & 2 & 4 \end{bmatrix}} = \underset{C}{\begin{bmatrix} 10 & 4 & 7 \\ 15 & 8 & 19 \end{bmatrix}}$$

Entonces

$$L_1^C = a_{11}L_1^B + a_{12}L_2^B + a_{13}L_3^B$$

$$L_2^C = a_{21}L_1^B + a_{22}L_2^B + a_{33}L_3^B$$

o sea

$$(10, 4, 7) = 2(2, 1, 3) + (-3)(-1, 0, 1) + (3, 2, 4)$$

$$(15, 8, 19) = 4(2, 1, 3) + (-1)(-1, 0, 1) + 2(3, 2, 4)$$

como se puede comprobar directamente.

Véase ahora el siguiente lema que se usará en la demostración del teorema 7.1.

LEMA Sea V un espacio vectorial y sean $S_1 = \{v_1, v_2, \ldots, v_r\}$ y $S_2 = \{u_1, u_2, \ldots, u_t\}$ dos conjuntos de vectores de V. Llámese W al espacio generado por los vectores $v_1, v_2, \ldots, v_r$. Si cada vector u_j, $1 \leq j \leq t$ se puede escribir como combinación lineal de los vectores v_i, $1 \leq i \leq r$, entonces

$$\mathcal{L}(u_1, u_2, \ldots, u_t) \subseteq W$$

DEMOSTRACIÓN Ciertamente para cada j, $1 \leq j \leq t$ se tiene $u_j \in W$. Si $\alpha_1, \alpha_2, \ldots, \alpha_t$ son escalares arbitrarios se tiene también que

$$\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_t u_t \in W$$

pues W es un subespacio (cerrado bajo la suma y el producto por escalares). Esto

muestra entonces que*

$$\mathcal{L}(u_1, u_2, \ldots, u_t) \subseteq W$$

Q.E.D.

TEOREMA 7.1 Sea la matriz $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$, denótese por A' a la forma escalonada reducida de A. Entonces el espacio línea de A es igual al espacio línea de A'.

DEMOSTRACIÓN Se debe mostrar que

$$EL^A = \mathcal{L}(L_1^A, L_2^A, \ldots, L_m^A) = \mathcal{L}(L_1^{A'}, L_2^{A'}, \ldots, L_m^{A'}) = EL^{A'}$$

Según el teorema 4.9 del capítulo 1, existe una matriz inversible Q de orden m tal que $A = QA'$.

Entonces (por la observación previa al lema) cada vector L_i^A es una combinación de los vectores $L_i^{A'}$. Por el lema anterior (con $S_1 = \{L_1^{A'}, L_2^{A'}, \ldots, L_m^{A'}\}$ y $S_2 = \{L_1^A, L_2^A, \ldots, L_m^A\}$) se concluye entonces que

$$EL^A \subseteq EL^{A'}$$

Pero siendo Q inversible, se puede escribir $A' = Q^{-1}A$, de modo que por un argumento análogo al anterior se obtiene

$$EL^{A'} \subseteq EL^A$$

o sea que $EL^A = EL^{A'}$, tal y como se quería.

Q.E.D.

EJEMPLO 2 Por ejemplo, si se tiene la matriz

$$A = \begin{bmatrix} 1 & 3 & 4 & 2 \\ 2 & 1 & 3 & -1 \\ 3 & 4 & 7 & 1 \end{bmatrix}$$

su forma escalonada reducida es:

*Otra demostración: Tómese $x \in \mathcal{L}(u_1, u_2, \ldots, u_t)$, entonces x se escribe como combinación lineal de los vectores $u_1, u_2, \ldots, u_t$. Pero por hipótesis, cada uno de estos vectores es una combinación lineal de los vectores $v_1, v_2, \ldots, v_r$. Poniendo todo esto junto, se ve que x puede escribirse como combinación lineal de los vectores $v_1, v_2, \ldots, v_r$. Entonces $x \in W$.

$$A' = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

El teorema anterior asegura que los dos subespacios de $\mathbf{R}^4$

$$EL^A = \mathcal{L}((1, 3, 4, 2), (2, 1, 3, -1), (3, 4, 7, 1))$$
$$EL^{A'} = \mathcal{L}((1, 0, 1, -1), (0, 1, 1, 1), (0, 0, 0, 0))$$
$$= \mathcal{L}((1, 0, 1, -1), (0, 1, 1, 1))$$

son iguales.

Además, observe que los vectores línea no nulos de la matriz A' son linealmente independientes pues

$$c_1(1, 0, 1, -1) + c_2(0, 1, 1, 1) = (0, 0, 0, 0)$$

implica que

$$\left.\begin{aligned} c_1 \quad &= 0 \\ c_2 &= 0 \\ c_1 + c_2 &= 0 \\ -c_1 + c_2 &= 0 \end{aligned}\right\} \Rightarrow \quad c_1 = c_2 = 0$$

Por lo tanto, los vectores $v_1 = (1, 0, 0, -1)$ y $v_2 = (0, 1, 1, 1)$ forman una base del espacio línea de A', y entonces, del espacio línea de A. Esto es un hecho general que se enuncia y demuestra a continuación.

TEOREMA 7.2

> Los vectores línea distintos de cero en la forma escalonada reducida de una matriz A constituyen una base para su espacio línea.

DEMOSTRACIÓN Sean $L_1^{A'}, L_2^{A'}, \ldots, L_r^{A'}$ los vectores línea distintos de cero en la forma escalonada reducida de A. Por el teorema 7.1, estos vectores generan al espacio línea de A.*

Considérese la combinación lineal

$$c_1 L_1^{A'} + \ldots + c_r L_r^{A'} = 0 \qquad (7.1)$$

Se demostrará que $c_1 = 0$.

Recuérdese que en A' el primer elemento distinto de cero en su primera línea es 1. Éste se encuentra, por decir, en la columna j_1. Además, en las líneas restantes los elementos correspondientes de la j_1-ésima columna (en donde se encuentra el primer elemento distinto de cero de la primera línea) son todos cero.

*No se están considerando los vectores línea que son cero en el espacio línea de A', pues es claro que éstos no alteran tal espacio.

En otras palabras, la primera coordenada distinta de cero en $L_1^{A'}$ es la j_1-ésima coordenada que es igual a 1, y en los vectores restantes $L_2^{A'}, \ldots, L_r^{A'}$ la j_1-ésima coordenada es cero.

Fíjese la atención en la j_1-ésima coordenada del vector de la izquierda en la expresión (7.1).

$$c_1(\ldots, \underbrace{1}_{j_1}, \ldots) + c_2(\ldots, \underbrace{0}_{j_1}, \ldots) + c_r(\ldots, \underbrace{0}_{j_1}, \ldots) = 0$$

se ve entonces que $c_1 = 0$.

Queda así (7.1) como

$$c_2 L_2^{A'} + \ldots + c_r L_r^{A'} = 0$$

Repítase el argumento anterior para mostrar que $c_2 = 0$.

De la misma manera se muestra que $c_3 = \ldots = c_r = 0$, y entonces los vectores $L_1^{A'}, \ldots, L_r^{A'}$ son *linealmente independientes* y por tanto constituyen una base del espacio EL^A.

Q.E.D.

El teorema anterior proporciona el fundamento del siguiente método para obtener una base del espacio generado por un conjunto de k vectores $v_1, v_2, \ldots, v_k$ en $\mathbf{R}^n$.

1) Fórmese la matriz A que tenga por vectores línea a los vectores dados $v_1, v_2, \ldots, v_k$ (A es entonces una matriz de orden $k \times n$).
2) Llévese la matriz A a su forma escalonada reducida.
3) Los vectores distintos de cero en la forma escalonada reducida de la matriz A constituyen una base del espacio línea de A (teorema anterior). Ésta será entonces una base para el espacio generado por los vectores $v_1, v_2, \ldots, v_k$.

EJEMPLO 3 Por ejemplo, suponga que se quiere determinar una base para el subespacio de $\mathbf{R}^5$ generado por los vectores $v_1 = (1, 2, -1, 4, 3)$, $v_2 = (3, 1, 2, 2, 4)$, $v_3 = (2, 1, 1, -1, 0)$ y $v_4 = (2, -4, 6, -11, -5)$.

Forme entonces la matriz

$$A = \begin{bmatrix} 1 & 2 & -1 & 4 & 3 \\ 3 & 1 & 2 & 2 & 4 \\ 2 & 1 & 1 & -1 & 0 \\ 2 & -4 & 6 & -11 & -5 \end{bmatrix}$$

Al llevar ésta a la forma escalonada reducida se obtiene:

$$A' = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

de modo que los vectores $u_1 = (1, 0, 1, 0, 1)$, $u_2 = (0, 1, -1, 0, -1)$ y $u_3 = (0, 0, 0, 1, 1)$ constituyen la base procurada. (El espacio que generan los vectores v_1, v_2, v_3, v_4 es entonces de dimensión 3.)

El método anterior se puede usar en cualquier espacio de dimensión finita, (dígase n), por medio de la identificación, a través de una base concreta β de V, de los vectores $W \in V$ con sus vectores de coordenadas.

Esto se puede hacer, pues, como se vio en la sección anterior, V es un espacio isomorfo a $\mathbf{R}^n$ y se cumple: β es una base de V si, y sólo si $f(\beta)$ es una base de $\mathbf{R}^n$ en donde: $f\colon V \to \mathbf{R}^n$ es un isomorfismo entre V y $\mathbf{R}^n$ (hecho que se demostrará en el próximo capítulo; véase teorema 6.4).

EJEMPLO 4 Por ejemplo, considérese en el espacio vectorial P_3 los vectores

$$p_1 = 2 + x + 3x^2 - x^3$$
$$p_2 = 4 + 5x + 3x^2 - x^3$$
$$p_3 = -1 + x - 3x^2 + x^3$$
$$p_4 = 1 + 5x - 3x^2 + x^3$$

Se quiere encontrar una base para el subespacio de P_3 generado por estos cuatro vectores.

Sea β la base canónica de P_3, esto es, $\beta = \{1, x, x^2, x^3\}$

Entonces

$$(p_1)_\beta = (2, 1, 3, -1) \qquad (p_3)_\beta = (-1, 1, -3, 1)$$
$$(p_2)_\beta = (4, 5, 3, -1) \qquad (p_4)_\beta = (1, 5, -3, 1)$$

Se forma la matriz

$$A = \begin{bmatrix} 2 & 1 & 3 & -1 \\ 4 & 5 & 3 & -1 \\ -1 & 1 & -3 & 1 \\ 1 & 5 & -3 & 1 \end{bmatrix}$$

Lleve esta matriz a su forma escalonada reducida, obteniendo

$$A' = \begin{bmatrix} 1 & 0 & 2 & -2/3 \\ 0 & 1 & -1 & 1/3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Entonces, los polinomios

$$q_1 = 1 + 2x^2 - \frac{2}{3}x^3$$

$$q_2 = x - x^2 + \frac{1}{3}x^3$$

forman una base del espacio generado por p_1, p_2, p_3, p_4, el cual es entonces de dimensión 2.

Para terminar, vea un ejemplo en el que intervienen varias ideas que se han estudiado a lo largo del presente capítulo.

EJEMPLO 5 Considere los 3 vectores de $\mathbf{R}^4$

$$v_1 = (1, -1, 0, 2)$$
$$v_2 = (2, 1, 1, 3)$$
$$v_3 = (-1, 7, 2, -4)$$

Estos vectores generan un cierto subespacio de $\mathbf{R}^4$ que se llamará W.

Por otra parte, se ha visto que las soluciones de un sistema homogéneo de ecuaciones lineales forman un espacio vectorial. Más concretamente, si se tiene un sistema homogéneo de ecuaciones en las incógnitas x_1, x_2, x_3, x_4, el espacio solución será (según se discutió en la sección 3) un subespacio de $\mathbf{R}^4$.

Uno se pregunta entonces por un sistema homogéneo de ecuaciones lineales (con 4 incógnitas) cuyo espacio solución sea precisamente W.

El primer paso que se tiene que dar para resolver este problema es describir explícitamente el subespacio W. Esto significa, por ejemplo, dar una base de él.

Según lo visto en esta sección, para hallar una base de W se forma la matriz

$$A = \begin{bmatrix} 1 & -1 & 0 & 2 \\ 2 & 1 & 1 & 3 \\ -1 & 7 & 2 & -4 \end{bmatrix}$$

y se lleva a su forma escalonada reducida.

Ésta es

$$A' = \begin{bmatrix} 1 & 0 & 1/3 & 5/3 \\ 0 & 1 & 1/3 & -1/3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Por lo tanto, una base de W está dada por los vectores

$$u_1 = \left(1, 0, \frac{1}{3}, \frac{5}{3}\right)$$

$$u_2 = \left(0, 1, \frac{1}{3}, -\frac{1}{3}\right)$$

Se tiene entonces que un vector $(x_1, x_2, x_3, x_4) \in \mathbf{R}^4$ se encuentra en W si, y sólo si existen escalares c_1 y c_2 tales que

$$(x_1, x_2, x_3, x_4) = c_1\left(1, 0, \frac{1}{3}, \frac{5}{3}\right) + c_2\left(0, 1, \frac{1}{3}, -\frac{1}{3}\right)$$

o lo que es lo mismo

$$\begin{aligned} c_1 &= x_1 \\ c_2 &= x_2 \\ \frac{1}{3}c_1 + \frac{1}{3}c_2 &= x_3 \\ \frac{5}{3}c_1 - \frac{1}{3}c_2 &= x_4 \end{aligned}$$

Éste es un sistema no homogéneo de 4 ecuaciones en las incógnitas c_1 y c_2. Uno se pregunta entonces qué relación deben guardar x_1, x_2, x_3, x_4 para que este sistema tenga solución para c_1 y c_2 (y por tanto para que $(x_1, x_2, x_3, x_4) \in W$). Con el método de eliminación Gaussiana se obtiene

$$\begin{bmatrix} 1 & 0 & x_1 \\ 0 & 1 & x_2 \\ 1/3 & 1/3 & x_3 \\ 5/3 & -1/3 & x_4 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & x_1 \\ 0 & 1 & x_2 \\ 0 & 0 & -x_1/3 - x_2/3 + x_3 \\ 0 & 0 & -5x_1/3 + x_2/3 + x_4 \end{bmatrix}$$

Entonces $(x_1, x_2, x_3, x_4) \in W$ si, y sólo si las dos últimas líneas de la matriz de la derecha constan sólo de ceros. Es decir, si

$$x_1 + x_2 - 3x_3 = 0$$
$$5x_1 - x_2 - 3x_4 = 0$$

Por otra parte, éste es un sistema homogéneo de dos ecuaciones lineales en las incógnitas x_1, x_2, x_3, x_4. Se afirma que (x_1, x_2, x_3, x_4) es solución de este sistema, lo que equivale a afirmar que $(x_1, x_2, x_3, x_4) \in W$. Por lo tanto, es el sistema requerido, cuyo espacio solución es precisamente W.

EJERCICIOS (SECCIÓN 7, CAPÍTULO 3)

1. Verifique el teorema 7.1 con la matriz

$$A = \begin{bmatrix} 1 & 2 & -1 & 0 & 7 \\ 3 & 4 & 1 & 1 & 1 \\ 0 & -2 & 4 & 1 & -20 \\ 7 & 10 & 1 & 2 & 4 \end{bmatrix}$$

2. Sea A una matriz inversible de orden n, describa el espacio línea de A.
3. Encuentre una base para el subespacio de $\mathbf{R}^4$ generado por los vectores $(1, 3, -1, 2)$, $(-1, 2, -2, 2)$ y $(2, 1, 1, 0)$.
4. Determine la dimensión de cada uno de los siguientes subespacios de $\mathbf{R}^4$
   a) $\mathcal{L}((1, 2, 1, 1), (3, -1, 2, 1))$
   b) $\mathcal{L}((1, 1, 1, 1), (2, 2, 2, 2))$
   c) $\mathcal{L}((3, 1, 2, 4))$
   d) $\mathcal{L}((1, 0, -1, 2), (6, 1, 1, 2), (4, 1, 3, -2), (8, 1, -1, 6))$
   e) $\mathcal{L}((0, 3, -1, 1), (2, 0, 0, 3), (4, 1, 7, 2), (3, 2, 3, 3))$
   f) $\mathcal{L}((1, 0, 0, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1))$
5. Encuentre una base para cada uno de los siguientes subespacios de $M_{2\times 2}$
   a) $\mathcal{L}\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\right)$
   b) $\mathcal{L}\left(\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 3 & 3 \end{bmatrix}\right)$
   c) $\mathcal{L}\left(\begin{bmatrix} 1 & -1 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 2 \\ 6 & 2 \end{bmatrix}\right)$
   d) $\mathcal{L}\left(\begin{bmatrix} -1 & 1 \\ -3 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 3 \\ 9 & 2 \end{bmatrix}, \begin{bmatrix} -1 & 4 \\ 12 & 1 \end{bmatrix}\right)$
6. Encuentre una base para cada uno de los siguientes subespacios de P_3
   a) $\mathcal{L}(-1 + x - x^2, 3x + x^2)$
   b) $\mathcal{L}(2 + x + x^2, -1 + 5x + 6x^2, -5 + 3x + 4x^2)$
   c) $\mathcal{L}(1 - x - x^2, 4 + 5x + x^2, 11 + 7x - x^2, 7 + 2x - 2x^2)$
   d) $\mathcal{L}(-1 + 2x, 5 + x - x^2, 12 + 9x - 3x^2)$

② 7. Sean W_1 y W_2 los subespacios de $\mathbf{R}^n$ dados por

$$W_1 = \mathcal{L}(v_1, v_2, \ldots, v_k)$$
$$W_2 = \mathcal{L}(u_1, u_2, \ldots, u_k)$$

Demuestre que una base del espacio $W_1 + W_2$ está constituida por los vectores diferentes de cero en la forma escalonada reducida de la matriz A de orden $(k + r) \times n$ cuyos vectores línea son $v_1, \ldots, v_k, u_1, \ldots, u_r$.

8. Sean W_1 y W_2 los subespacios de $\mathbf{R}^4$ dados por

$$W_1 = \mathcal{L}((2, 3, -1, 1), (8, 12, -9, 8), (2, 3, 9, -7))$$

$$W_2 = \mathcal{L}((4, 6, 3, -2), (6, 9, -10, 9), (2, 3, -28, 23))$$

Encuentre una base para los subespacios $W_1 \cap W_2$ y $W_1 + W_2$.

9. Repita el ejercicio anterior con los subespacios

$$W_1 = \mathcal{L}((1, 1, 1, 1), (1, -1, 1, -1), (1, 3, 1, 3))$$

$$W_2 = \mathcal{L}((1, 2, 0, 2), (1, 2, 1, 2), (3, 1, 3, 1))$$

② 10. Encuentre un sistema homogéneo de ecuaciones lineales cuyo espacio solución sea:

a) La recta $x = 2t, y = t, z = 3t, t \in \mathbf{R}$.

b) El plano $5x + 2y - 3z = 0$.

c) El subespacio $\mathbf{R}^4$ generado por $(2, 1, 3, 0), (1, -1, 1, 1), (7, -1, 9, 3)$

d) El subespacio de $\mathbf{R}^5$ generado por $(1, -1, 2, 1, 0), (4, 0, 2, 0, -1), (3, 1, 0, -1, -1), (4, 6, 2, 0, -1)$.

CAPÍTULO CUATRO

Transformaciones lineales

En la introducción se estableció el concepto de *función* de un conjunto A a un conjunto B, $f: A \to B$, como una regla que a cada elemento a del conjunto A (el dominio de la función) le hace corresponder un único elemento $b = f(a)$ del conjunto B (el codominio de la función).

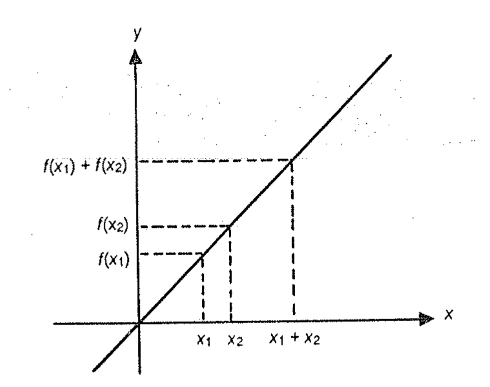
En la experiencia previa en cursos de matemáticas (por ejemplo en los cursos de cálculo) se ha trabajado con funciones cuyo codominio es el conjunto de números reales $\mathbf{R}$ y cuyo dominio es algún subconjunto de $\mathbf{R}$. Es decir, funciones de la forma $f: I \subseteq \mathbf{R} \to \mathbf{R}$. Un tipo especial de estas funciones son las llamadas "funciones lineales", las cuales son $f: \mathbf{R} \to \mathbf{R}$, $f(x) = mx + b$, en donde m y b son números reales determinados. Estas funciones, como se sabe, representan geométricamente (en el plano xy) rectas de pendiente m y ordenada al origen b.

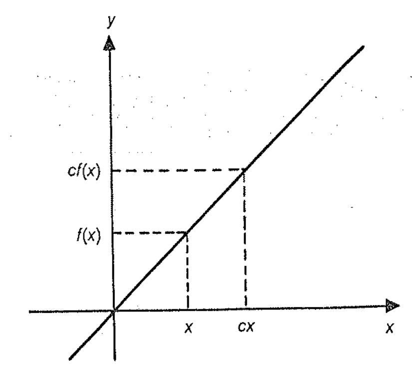
Se desea resaltar ahora un par de propiedades que poseen algunas funciones lineales: considérese la función lineal $f: \mathbf{R} \to \mathbf{R}$, $f(x) = mx$. Su gráfica es entonces una recta de pendiente m que pasa por el origen (su ordenada al origen es cero). Obsérvese que se tiene

1) $f(x_1 + x_2) = m(x_1 + x_2) = mx_1 + mx_2 = f(x_1) + f(x_2)$

2) $f(cx) = m(cx) = c(mx) = cf(x)$

Es decir, la función f transforma "sumas en sumas" y "productos por escalares en productos por escalares". Geométricamente esto se ve como





Desde la perspectiva de los cursos de cálculo, estas propiedades resultan ser poco interesantes, pues la *mayoría* de las funciones que en tales cursos aparecen *no* cumplen con ellas (por ejemplo, si f: **R** → **R** es $f(x)$ = sen x, no es cierto que $f(x_1 + x_2) = \text{sen}(x_1 + x_2) = \text{sen}\, x_1 + \text{sen}\, x_2 = f(x_1) + f(x_2)$ y $f(cx) = \text{sen}\, cx = c\, \text{sen}\, x = cf(x)$). Sin embargo, desde el punto de vista del álgebra lineal, éstas serían (las funciones $f(x) = mx$) *las funciones importantes a estudiar.*

En este capítulo es de interés la generalización (la abstracción) algebraica de este tipo de funciones; estas funciones tendrán por dominio y codominio a *espacios vectoriales* (de los que un caso particular es **R**, como en el ejemplo) y cumplirán los propiedades 1) y 2) citadas anteriormente: transformarán sumas (de elementos —de vectores— en su dominio) en sumas (de las correspondientes imágenes en su codominio) y producto por escalares (cv, v un vector del dominio) en producto por escalares ($cf(v)$, $f(v)$ la imagen de v). A estas funciones se les llamará *transformaciones lineales.*

Las transformaciones lineales son el tipo de funciones más importantes estudiadas en el álgebra lineal. Éstas representan el prototipo de "buen comportamiento" (¡desde el punto de vista del álgebra lineal!) de una función. El capítulo anterior se ha dedicado entonces a estudiar la estructura algebraica del dominio y codominio de estas funciones.

Este capítulo, al igual que el anterior, constituye uno de los temas centrales en el estudio del álgebra lineal.

1. DEFINICIÓN Y EJEMPLOS

Sean V y U dos espacios vectoriales. Se dice que la función

$$T: V \to U$$

es una *transformación lineal* si

1) T preserva sumas

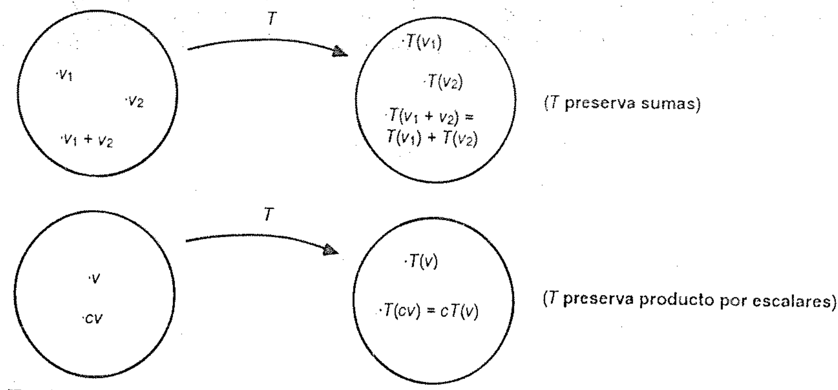
$$T(v_1 + v_2) = T(v_1) + T(v_2) \qquad \forall v_1, v_2 \in V$$

2) T preserva producto por escalares

$$T(cv) = cT(v) \qquad \forall v \in V, \qquad \forall c \in \mathbf{R}$$

En el caso particular en que $V = U$ se dice que T es un *operador lineal.* Una transformación lineal es, pues, una función entre dos espacios vectoriales "que preserva las operaciones de espacio vectorial", en el sentido de que: 1) la imagen de una suma de vectores en su dominio es la suma de las imágenes de cada uno de los vectores, y 2) la imagen del producto de un vector en su dominio por un escalar es el producto de la imagen del vector por el escalar.

Esquemáticamente



En la siguiente "observación preliminar" se obtienen algunas consecuencias inmediatas de la definición de transformación lineal.

OBSERVACIÓN PRELIMINAR. Sea T: $V \to U$ una transformación lineal. Entonces

1) $T(0) = 0$ (la imagen del vector cero en V es el vector cero en U).
2) $T(-v) = -T(v)$ (la imagen del inverso aditivo del vector $v \in V$ es el inverso aditivo del vector $T(v) \in U$).
3) $T\left(\sum_{i=1}^{n} c_i v_i\right) = \sum_{i=1}^{n} c_i T(v_i)$, donde $c_i, c_2, \ldots, c_n \in \mathbf{R}$, $v_i, v_2, \ldots, v_n \in V$

(T preserva combinaciones lineales de vectores en V).

En efecto:

1) $T(0) = T(0 + 0) = T(0) + T(0)$

En vista de la unicidad del vector cero en U, se concluye que $T(0) = 0$.

2) $T(-v) = T((-1)v) = (-1)T(v) = -T(v)$.
3) Este resultado se obtiene por una aplicación repetida de la definición de transformación lineal. Se deja al lector que dé un argumento detallado que pruebe esta propiedad (véase observación en la sección 1 del capítulo 3 sobre cómo se define una suma de n vectores en un espacio vectorial).

El resto de la presente sección se dedicará a ver varios ejemplos de transformaciones lineales. Algunos de los ejemplos que se darán serán "triviales" (pero importantes de considerar), algunos otros serán (no triviales y) analíticos y también se verá un ejemplo puramente geométrico de una transformación lineal.

EJEMPLO 1 El primer ejemplo que se verá (uno de los ejemplos triviales) es la *transformación cero*: sea T: $V \to U$ la transformación tal que $T(v) = 0\ \forall\, v \in V$. Entonces T es una transformación lineal pues

$$T(v_1 + v_2) = 0 = 0 + 0 = T(v_1) + T(v_2)$$

y

$$T(cv) = 0 = c \cdot 0 = cT(v)$$

Se denotará a esta transformación por 0.

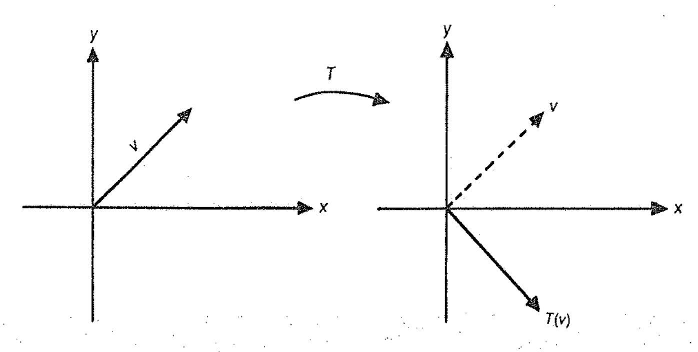
EJEMPLO 2 Otro ejemplo (también trivial como el anterior) está dado por la *transformación identidad* $T: V \to V$ definida por $T(v) = v \; \forall v \in V$ (obsérvese que T es un operador lineal). T es una transformación lineal pues

$$T(v_1 + v_2) = v_1 + v_2 = T(v_1) + T(v_2)$$

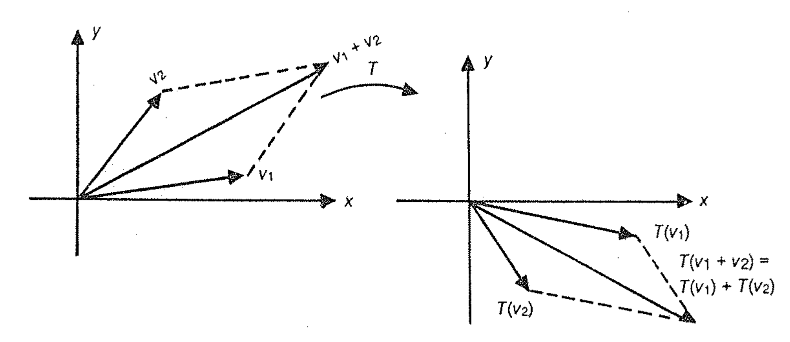
$$T(cv) = cv = cT(v)$$

Se denotará esta transformación por Id (o más explícitamente como Id_V).

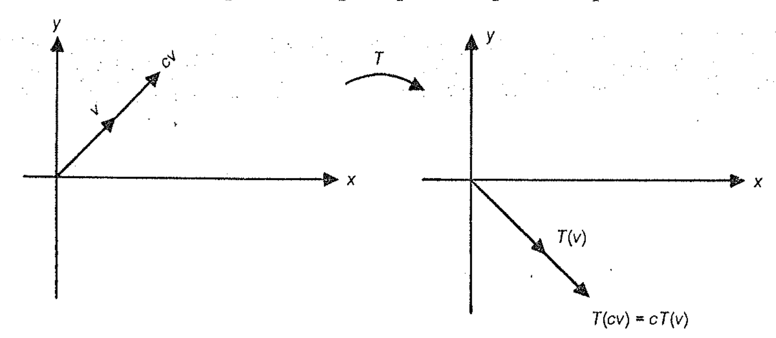
EJEMPLO 3 Véase ahora un ejemplo geométrico. Considérese de transformación $T: \mathbf{R}^2 \to \mathbf{R}^2$ definida de la siguiente manera: para $v = (x, y) \in \mathbf{R}^2$, $T(v)$ es el vector de $\mathbf{R}^2$ que geométricamente representa la reflexión del vector v respecto del eje x.



Al ver la siguiente figura uno se puede convencer de que T preserva sumas

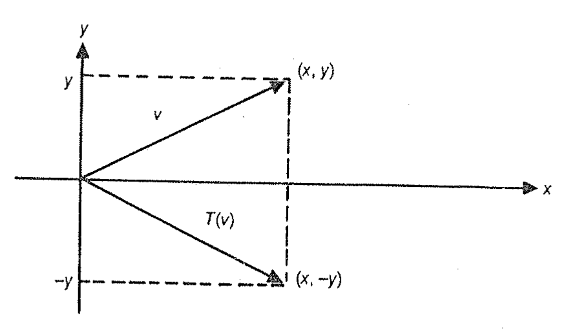


De una manera similar se puede ver que T preserva producto por escalares



Se puede hacer una descripción analítica de la transformación anterior y verificar "analíticamente" que se trata de una transformación lineal.

Obsérvese que



de modo que nuestra descripción analítica de T es:

$$T: \mathbf{R}^2 \to \mathbf{R}^2$$

$$T(x, y) = (x, -y)$$

Tómese $v_1 = (x, y)$, $v_2 = (x', y')$ y verifíquese que $T(v_1 + v_2) = T(v_1) + T(v_2)$

$$\begin{aligned} T(v_1 + v_2) &= T((x, y) + (x', y')) = T(x + x', y + y') \\ &= (x + x', -(y + y')) = (x + x', -y - y') = (x, -y) + (x', -y') \\ &= T(v_1) + T(v_2) \end{aligned}$$

Similarmente, si $c \in \mathbf{R}$

$$\begin{aligned} T(cv) &= T(c(x, y)) = T(cx, cy) = (cx, -cy) = (cx, c(-y)) \\ &= c(x, -y) = cT(v) \end{aligned}$$

EJEMPLO 4 Considere ahora la transformación $T: \mathbf{R}^3 \to \mathbf{R}^2$ dada por

$$T(x, y, z) = (2x + 3y - 4z, x + 2y + z)$$

Sean $v_1 = (x, y, z)$, $v_2 = (x', y', z')$ dos vectores de $\mathbf{R}^3$. Se tiene

$$\begin{aligned} T(v_1 + v_2) &= T((x, y, z) + (x', y', z')) = T(x + x', y + y', z + z') \\ &= (2(x + x') + 3(y + y') - 4(z + z'), x + x' + 2(y + y') + z + z') \\ &= (2x + 2x' + 3y + 3y' - 4z - 4z', x + x' + 2y + 2y' + z + z') \\ &= (2x + 3y - 4z + 2x' + 3y' - 4z', x + 2y + z + x' + 2y' + z') \\ &= (2x + 3y - 4z, x + 2y + z) + (2x' + 3y' - 4z', x' + 2y' + z') \\ &= T(x, y, z) + T(x', y', z') = T(v_1) + T(v_2) \end{aligned}$$

Similarmente, si $c \in \mathbf{R}$

$$\begin{aligned} T(cv) &= T(c(x, y, z)) = T(cx, cy, cz) \\ &= (2cx + 3cy - 4cz,\ cx + 2cy + cz) \\ &= c(2x + 3y - 4z, x + 2y + z) = cT(x, y, z) = cT(v) \end{aligned}$$

lo que muestra que T es lineal.

Recapacítese un poco sobre este último ejemplo. Considérese la matriz A de orden 2×3.

$$A = \begin{bmatrix} 2 & 3 & -4 \\ 1 & 2 & 1 \end{bmatrix} \tag{1.1}$$

Al escribir los vectores de $\mathbf{R}^n$ como matrices $n \times 1$, se tiene para $v = (x, y, z)$

$$Av = \begin{bmatrix} 2 & 3 & -4 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + 3y - 4z \\ x + 2y + z \end{bmatrix}$$

Obsérvese entonces que la multiplicación de A por el vector $v \in \mathbf{R}^3$ produce precisamente el vector $T(v) \in \mathbf{R}^2$. Por tanto, otra descripción que se puede dar de la transformación del ejemplo anterior es como T: $\mathbf{R}^3 \to \mathbf{R}^2$, $T(v) = Av$, donde A es la matriz (1.1) y se entiende que en la expresión Av se está considerando al vector $v \in \mathbf{R}^3$ como una matriz 3×1 (la identificación del isomorfismo natural $\mathbf{R}^n \cong M_{n \times 1}$).

Obsérvese también que al usar esta "descripción matricial" de T es muy fácil verificar que se trata de una transformación lineal. En efecto, al usar lo que se sabe del álgebra de matrices

$$\begin{aligned} T(v_1 + v_2) &= A(v_1 + v_2) = Av_1 + Av_2 = T(v_1) + T(v_2) \\ T(cv) &= A(cv) = cAv = cT(v) \end{aligned} \tag{1.2}$$

lo que muestra como antes, con un ahorro considerable de espacio y de tinta, que T es lineal.

EJEMPLO 5 Más generalmente, considérese la transformación T: $\mathbf{R}^n \to \mathbf{R}^m$ dada por

$$T(v) = Av$$

donde A es una matriz $m \times n$ ($v \in \mathbf{R}^n$, o sea, una matriz $n \times 1$, de modo que Av es una matriz $m \times 1$, o bien, $Av \in \mathbf{R}^m$, la imagen de v bajo T).

T es una transformación lineal puesto que siguen siendo válidas en este caso las fórmulas (1.2).

Éste es un ejemplo importante pues se verá después (sección 3) que siendo T una transformación lineal del espacio V de dimensión n, al espacio U de dimensión m, se podrá asociar a esta transformación (por medio de bases concretas de V y U) una matriz A de orden $m \times n$, de modo que T "puede verse" como la transformación del ejemplo anterior.

EJEMPLO 6 Considere ahora la transformación T: $P_n \to P_n$ dada por

$$T(p) = p'$$

donde p' denota la derivada del polinomio $p \in P_n$.

Más explícitamente

$$T(a_0 + a_1x + \ldots + a_nx^n) = a_1 + 2a_2x + \ldots + na_nx^{n-1}$$

T es una transformación lineal pues:

1) $T(p_1 + p_2) = (p_1 + p_2)' = p_1' + p_2' = T(p_1) + T(p_2)$. (La derivada de una suma de funciones es la suma de las derivadas de cada una de las funciones.)
2) $T(cp) = (cp)' = cp' = cT(p)$. (La derivada de una constante por una función es la constante por la derivada de la función.)

EJEMPLO 7 Sea ahora $V = C([a, b])$, el espacio de las funciones continuas en el intervalo $[a, b]$. Defínase T: $V \to \mathbf{R}$ como

$$T(f) = \int_a^b f(x)dx$$

Que T es una transformación lineal se puede verificar (como en el ejemplo anterior) usando los conocimientos del curso de cálculo.

1) $T(f_1 + f_2) = \int_a^b (f_1 + f_2)(x)dx = \int_a^b (f_1(x) + f_2(x))dx$

$$= \int_a^b f_1(x)dx + \int_a^b f_2(x)dx = T(f_1) + T(f_2)$$

2) $T(cf) = \int_a^b (cf)(x)dx = \int_a^b cf(x)dx = c\int_a^b f(x)dx = cT(f)$

en donde se usa que la integral de una suma de funciones es la suma de las integrales de cada una de las funciones y la integral de una constante por una función es la constante por la integral de la función ("las constantes salen del signo de integral").

Véase ahora un ejemplo de una transformación entre espacios vectorial que no es lineal.

EJEMPLO 8 Sea $T: \mathbf{R}^2 \to \mathbf{R}$ la transformación dada por

$$T(x, y) = x + y + 1$$

Se tiene

$$T((x, y) + (x', y')) = T(x + x', y + y') = x + x' + y + y' + 1$$

Pero

$$T(x, y) + T(x', y') = x + y + 1 + x' + y' + 1 = x + x' + y + y' + 2$$

de modo que

$$T((x, y) + (x', y')) \neq T(x, y) + T(x', y')$$

por lo que T no es entonces una transformación lineal.

Para finalizar esta sección se verá un resultado básico sobre "construcción" de transformaciones lineales, el cual se usará posteriormente durante el desarrollo de la teoría que sobre transformaciones lineales se hará en este capítulo.

TEOREMA 1.1 Sean V y U dos espacios vectoriales, en donde V es de dimensión finita (dígase que $dim\ V = n$). Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V y sean $u_1, u_2, \ldots, u_n$ vectores cualesquiera de U. Existe una única transformación lineal $T: V \to U$ tal que $T(v_i) = u_i$, $i = 1, 2, \ldots, n$.

DEMOSTRACIÓN Para el vector $v \in V$, existen únicos escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ tales que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

Defínase $T: V \to U$ poniendo

$$T(v) = T(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n$$

Obsérvese que

$$T(v_i) = T(0 \cdot v_1 + \ldots + 1 \cdot v_i + \ldots + 0 v_n) = 0 \cdot u_1 + \ldots + 1 \cdot u_i + \ldots 0 \cdot u_n = u_i$$

de modo que esta transformación T posee la propiedad requerida en la conclusión del teorema.

Véase que T es una transformación lineal.

Sean $v, v' \in V$, dígase que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

$$v' = \alpha_1' v_1 + \alpha_2' v_2 + \ldots + \alpha_n' v_n$$

Entonces

$$v + v' = (\alpha_1 + \alpha_1')v_1 + (\alpha_2 + \alpha_2')v_2 + \ldots + (\alpha_n + \alpha_n')v_n$$

de modo que

$$T(v + v') = (\alpha_1 + \alpha_1')u_1 + (\alpha_2 + \alpha_2')u_2 + \ldots + (\alpha_n + \alpha_n')u_n$$
$$= (\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n) + (\alpha_1' u_1 + \alpha_2' u_2 + \ldots + \alpha_n' u_n)$$
$$= T(v) + T(v')$$

Si $c \in \mathbf{R}$

$$T(cv) = T(c\alpha_1 v_1 + c\alpha_2 v_2 + \ldots + c\alpha_n v_n) = c\alpha_1 u_1 + c\alpha_2 u_2 + \ldots + c\alpha_n u_n$$
$$= c(\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n) = cT(v)$$

lo que muestra entonces que T es lineal.

Para probar la unicidad de T, supongamos que $\tilde{T}: V \to U$ es una transformación lineal con la propiedad $\tilde{T}(v_i) = u_i$, $i = 1, 2, \ldots, n$. Entonces para cualquier $v \in V$, se tiene

$$\tilde{T}(v) = \tilde{T}(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) = \alpha_1 \tilde{T}(v_1) + \alpha_2 \tilde{T}(v_2) + \ldots + \alpha_n \tilde{T}(v_n)$$

↑ $\tilde{T}$ es lineal

$$= \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n = T(v)$$

es decir $\tilde{T}(v) = T(v)\ \forall\ v \in V$, lo que significa entonces que $\tilde{T} = T$.

Q.E.D.

Una consecuencia inmediata del teorema anterior se enuncia en el siguiente corolario, cuya demostración se deja como ejercicio para el lector.

COROLARIO Con la notación del teorema 1.1., si T_1 y T_2 son dos transformaciones lineales tales que $T_1(v_j) = T_2(v_j)$, $j = 1, 2, \ldots, n$ (T_1 y T_2 coinciden en los elementos de la base β de V), entonces $T_1 = T_2$.

EJERCICIOS (SECCIÓN 1, CAPÍTULO 4)

1. Determine cuáles de las siguientes transformaciones $T: \mathbf{R}^2 \to \mathbf{R}$ son lineales

   a) $T(x, y) = 2x + y$.
   b) $T(x, y) = x$.

c) $T(x, y) = y$.

d) $T(x, y) = 5x^2 + y$.

e) $T(x, y) = 3x - 8y + 1$.

f) $T(x, y) = 1$.

2. Determine cuáles de las siguientes transformaciones $T: \mathbf{R}^2 \to \mathbf{R}^2$ son lineales

a) $T(x, y) = (x + y, x - y)$

b) $T(x, y) = (5x, 8y)$

c) $T(x, y) = (1, x + y)$

d) $T(x, y) = (x + y, 0)$

e) $T(x, y) = (2x - y, 3x + 4y)$

f) $T(x, y) = (0, y)$

3. Determine cuáles de las siguientes transformaciones $T: \mathbf{R}^3 \to \mathbf{R}^3$ son lineales

a) $T(x, y, z) = (x, y, z)$

b) $T(x, y, z) = (x + y + z, 0, 0)$

c) $T(x, y, z) = (2x + y - z, 3x - y - 2z, x + 2y + 3z)$

d) $T(x, y, z) = (x, y + 1, z + 2)$

e) $T(x, y, z) = (0, 0, z)$

f) $T(x, y, z) = (x + y + z, 1, 1)$

4. Determine cuáles de las siguientes transformaciones $T: P_2 \to P_3$ son lineales

a) $T(a_0 + a_1x + a_2x^2) = a_0 + a_2x^2 + x^3$

b) $T(a_0 + a_1x + a_2x^2) = a_0 + 5 + x^2$

c) $T(a_0 + a_1x + a_2x^2) = (a_0 + a_1) + (a_0 + a_2)x + a_2x^2 + (a_1 + a_2)x^3$

d) $T(a_0 + a_1x + a_2x^2) = a_0 + a_0x + a_1x^2 + a_2x^3$

e) $T(a_0 + a_1x + a_2x^2) = (a_0 + a_1 + a_2)x^3$

f) $T(a_0 + a_1x + a_2x^2) = (a_0 - a_1) + (a_0 - a_2)x + (a_0 - a_1)x^2 + (a_0 - a_2)x^3$

5. Considere la transformación $T: M_{n\times n} \to M_{n\times n}$ dada por $T(A) = A^t$ (la transpuesta de la matriz A). Demuestre que T es una transformación lineal.

6. Sea B una matriz cuadrada de orden n. Considere la transformación $T: M_{n\times n} \to M_{n\times n}$ dada por $T(A) = AB - BA$ (B es una matriz fija de orden n). Demuestre que T es una transformación lineal. ¿Cuál es la imagen bajo esta transformación de una matriz de orden n que conmute con la matriz B? Dé al menos dos ejemplos distintos de matrices A para las cuales $T(A) = 0$.

7. a) Considere la transformación $T: M_{n\times n} \to \mathbf{R}$ dada por $T(A) = tr\ A$ (la traza de la matriz A). Demuestre que T es una transformación lineal.

b) ¿Es $T: M_{n\times n} \to \mathbf{R}$, $T(A) = det\ A$ una transformación lineal?

8. Demuestre que la transformación $T: \mathbf{R}^n \to \mathbf{R}$ dada por

$$T(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n$$

($a_1, a_2, \ldots, a_n$ son constantes dadas) es una transformación lineal.

9. Compruebe que la transformación $T: C^2(\mathbf{R}) \to C(\mathbf{R})$ dada por

$$T(f) = a_2f'' + a_1f' + a_0f$$

(a_0, a_1, a_2 son constantes dadas) es una transformación lineal.

10. Sean T_1 y T_2 dos transformaciones lineales del espacio vectorial V a $\mathbf{R}$. Demuestre que la transformación $S: V \to \mathbf{R}^2$ dada por $S(v) = (T_1(v), T_2(v))$ es una transformación lineal.

② 11. Sea V un espacio vectorial de dimensión 1 y sea $T: V \to V$ una transformación lineal. Demuestre que existe un escalar $c \in \mathbf{R}$ tal que $T(v) = cv\ \forall\ v \in V$.

① 12. Sea $T: V \to U$ una transformación lineal del espacio vectorial V al espacio vectorial U. Sean $v_1, v_2, \ldots, v_k$ vectores de V tales que $T(v_1), T(v_2), \ldots, T(v_k)$ son vectores linealmente independientes en U. Demuestre entonces que $v_1, v_2, \ldots, v_k$ son linealmente independientes. ¿Es válida la afirmación recíproca?

① 13. Sea $T: V \to U$ una transformación lineal del espacio vectorial V al espacio vectorial U. Sean $v_1, v_2, \ldots, v_k$ vectores de V tales que las imágenes $T(v_1), T(v_2), \ldots, T(v_k)$ generan al espacio U. Demuestre que T es sobreyectiva, esto es, pruebe que dado $u \in U$ existe $v \in V$ tal que $T(v) = u$.

14. Sea $T: V \to U$ una transformación lineal del espacio vectorial V al espacio vectorial U.

a) Si V_1 es un subespacio de V, demuestre que

$$T(V_1) = \{u \in U \mid u = T(v), v \in V_1\}$$

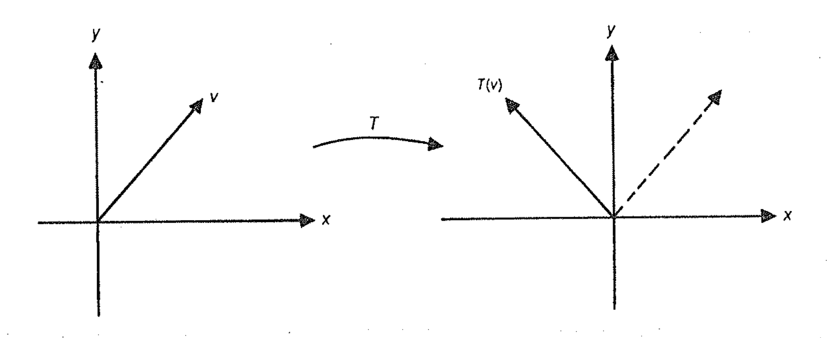
es un subespacio de U.

b) Si U_1 es un subespacio de U, demuestre que

$$T^{-1}(U_1) = \{v \in V \mid T(v) \in U_1\}$$

es un subespacio de V.

15. Considere la siguiente descripción geométrica de la transformación $T: \mathbf{R}^2 \to \mathbf{R}^2$: para un vector $v \in \mathbf{R}^2$, $T(v)$ es el vector de $\mathbf{R}^2$ reflexión respecto del eje y del vector v.



a) Demuestre que T es una transformación lineal.

b) ¿Cuál es la imagen bajo T de los vectores que se encuentran sobre el eje y?

16. Sea $T: \mathbf{R}^2 \to \mathbf{R}^2$ la transformación $T(x, y) = (x, 0)$.

a) Demuestre que T es una transformación lineal.

b) Dé una descripción geométrica de T.

c) ¿Cuál es la imagen bajo T de los vectores que se encuentran sobre el eje y?, ¿de los vectores que se encuentran sobre el eje x?

17. Sea T: $\mathbb{R}^2 \to \mathbb{R}^2$ la transformación $T(x, y) = (-x, -y)$.
    a) Demuestre que T es una transformación lineal.
    b) Dé una descripción geométrica de T.
18. Sea T: $\mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal tal que $T((1, 0)) = (2, 3)$ y $T((0, 1)) = (1, 1)$.
    a) Demuestre que T es única.
    b) Obtenga la imagen bajo T del vector $(3, 4)$.
    c) Obtenga la imagen bajo T de un vector cualquiera (x, y).
19. Sea T: $\mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal tal que $T((2, 5)) = (3, 1)$ y $T((1, 2)) = (0, 3)$. Calcule $T((3, 2))$.
20. Sea T: $\mathbb{R}^3 \to P_2$ la transformación lineal tal que

$$T((1, 1, 1)) = 1 - x + x^2$$
$$T((2, 0, 0)) = 3 + x - x^2$$
$$T((0, 4, 5)) = 2 + 3x - x^2$$

    Calcule $T((2, 4, -2))$.
21. Demuestre el corolario del teorema 1.1.
22. Sea T: $V \to W$ una transformación lineal del espacio vectorial V de dimensión finita al espacio vectorial W. Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V. Suponga que $T(v_1) = T(v_2) = \ldots = T(v_n) = 0$. Demuestre que T es la transformación cero.
23. Sea T: $V \to V$ un operador lineal en el espacio vectorial V de dimensión finita. Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V. Demuestre que si $T(v_i) = v_i$, $i = 1, 2, \ldots, n$, entonces T es el operador identidad en V.

2. NÚCLEO E IMAGEN

En la sección anterior se ha establecido la definición de transformación lineal y se ha ilustrado este concepto con varios ejemplos particulares. Se vio también un teorema básico, que es el primero de muchos resultados interesantes que se obtendrán a lo largo de este capítulo.

En esta sección se comienza formalmente el estudio de las transformaciones lineales. Existen dos conceptos fundamentales que intervienen en este estudio, que se refieren a dos subespacios (uno del dominio y otro del codominio), cuyo conocimiento reporta una gran cantidad de información sobre la "estructura interna" de la transformación lineal en consideración. Estos subespacios, cuyos nombres dan el título a esta sección, aparecerán continuamente en la teoría que se desarrollará en este capítulo. Ésta es, pues, la sección en la que se aborda de lleno el estudio de la teoría de transformaciones lineales.

DEFINICIÓN 2.1 Sea T: $V \to U$ una transformación lineal del espacio vectorial V al espacio vectorial U.

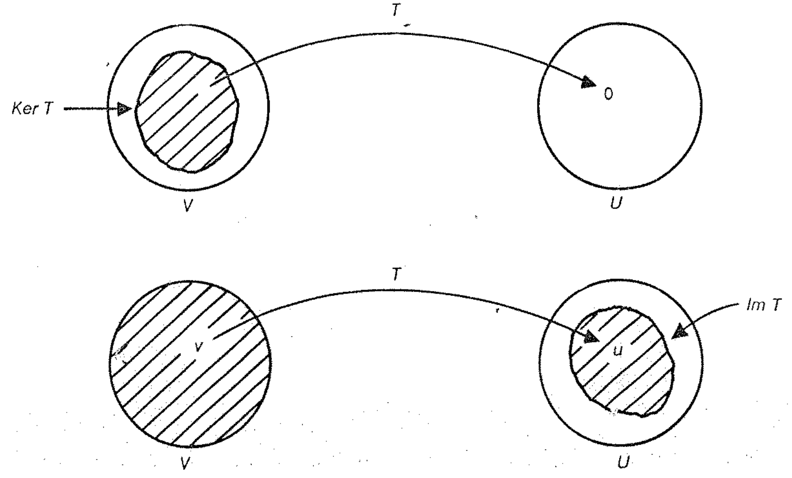
Defínase el *núcleo* de esta transformación, el cual se denotará por $Ker\ T$, como el subconjunto de V dado por

$$Ker\ T = \{v \in V \mid T(v) = 0\} = T^{-1}(0)$$

Se define también la imagen de la transformación T: $V \to U$ denotada por $Im\ T$ como el subconjunto de U dado por

$$Im\ T = \{u \in U \mid \exists\, v \in V, \text{ tal que } T(v) = u\}$$

Esquemáticamente estos conceptos se ven como



NOTA: Para una función f: $A \to B$ del conjunto A al conjunto B, se tiene definido el concepto de *rango* de la función como el subconjunto de B

$$\{b \in B \mid \exists\, a \in A \text{ tal que } f(a) = b\}$$

Con esta perspectiva lo que se ha definido como imagen de una transformación lineal T: $V \to U$ es precisamente el rango de la función T: $V \to U$. En el álgebra lineal, la palabra "rango" se destina para otro uso, el cual se verá en esta misma sección.

Para una transformación lineal T: $V \to U$, el núcleo y, la imagen no son sólo *subconjuntos* de V y U respectivamente. Resulta que éstos son *subespacios*. Esto es lo que dice el siguiente teorema:

TEOREMA 2.1 Sean V y U dos espacios vectoriales y T: $V \to U$ una transformación lineal. Entonces $ker\ T$ es un subespacio de V e $Im\ T$ es un subespacio de U.

DEMOSTRACIÓN Sean v y v' vectores de $Ker\ T$, entonces

$$\begin{aligned} T(v + v') &= T(v) + T(v') && \text{(pues } T \text{ es lineal)} \\ &= 0 + 0 && \text{(pues } v, v' \in Ker\ T) \\ &= 0 \end{aligned}$$

Entonces $v + v' \in Ker\ T$.

Similarmente, para $c \in \mathbf{R}$

$$T(cv) = cT(v) = c0 = 0$$

lo que prueba que $cv \in Ker\ T$, y entonces $Ker\ T$ es un subespacio de V.

Tómese ahora $u, u' \in Im\ T$. Esto significa entonces que existen $v, v' \in V$ tales que $T(v) = u$, $T(v') = u'$. Por lo tanto,

$$T(v + v') = T(v) + T(v') = u + u'$$

lo que implica que $u + u' \in Im\ T$ (pues existe $v + v' \in V$ tal que $T(v + v') = u + u'$).

También, si $c \in \mathbf{R}$

$$T(cv) = cT(v) = cu$$

lo que implica que $cu \in Im\ T$.

Entonces $Im\ T$ es un subespacio de U.

Q.E.D.

Véanse un par de ejemplos

EJEMPLO 1 Sea $T: \mathbf{R}^3 \to \mathbf{R}^3$ la transformación

$$T(x, y, z) = (x + 2y + z, 3x + y - 2z, -x - 7y - 6z)$$

Se verifica fácilmente que T es un operador lineal.

Se propone describir explícitamente su núcleo y su imagen (hallar, por ejemplo, una base para cada uno de estos subespacios).

Por definición

$$\begin{aligned} Ker\ T &= \{(x, y, z) \mid T(x, y, z) = 0\} \\ &= \{(x, y, z) \mid (x + 2y + z, 3x + y - 2z, -x - 7y - 6z) = 0\} \end{aligned}$$

Por tanto, el vector $(x, y, z) \in \mathbf{R}^3$ pertenece a $Ker\ T$ si, y sólo si

$$\begin{aligned} x + 2y + z &= 0 \\ 3x + y - 2z &= 0 \\ -x - 7y - 6z &= 0 \end{aligned}$$

Éste es un sistema homogéneo en las incógnitas x, y, z cuyo espacio solución es precisamente $Ker\ T$. Resuélvase por el método de eliminación Gaussiana

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 1 & -2 \\ -1 & -7 & -6 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

de donde $x = t$, $y = -t$, $z = t$, $t \in \mathbf{R}$, representa la solución del sistema.

Entonces se concluye que $(x, y, z) \in Ker\ T$ si, y sólo si

$$(x, y, z) = (t, -t, t) = t(1, -1, 1) \qquad t \in \mathbf{R}$$

El vector $v = (1, -1, 1)$ es entonces una base para el subespacio $Ker\ T$ (que geométricamente representa una recta que pasa por el origen $x = t$, $y = -t$, $z = t$, $t \in \mathbf{R}$).

Descríbase ahora $Im\ T$.

El vector $(b_1, b_2, b_3) \in Im\ T$ si, y sólo si existe $(x, y, z) \in \mathbf{R}^3$, tal que

$$\begin{aligned} x + 2y + z &= b_1 \\ 3x + y - 2z &= b_2 \\ -x - 7y - 6z &= b_3 \end{aligned}$$

Procédase por el método de eliminación Gaussiana para describir cómo tiene que ser el vector $(b_1, b_2, b_3) \in \mathbf{R}^3$ para que el sistema anterior posea solución

$$\begin{bmatrix} 1 & 2 & 1 & b_1 \\ 3 & 1 & -2 & b_2 \\ -1 & -7 & -6 & b_3 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -1 & -\frac{1}{5}b_1 + \frac{2}{5}b_2 \\ 0 & 1 & 1 & \frac{3}{5}b_1 - \frac{1}{5}b_2 \\ 0 & 0 & 0 & -\frac{4}{5}b_1 + \frac{1}{5}b_2 - \frac{1}{5}b_3 \end{bmatrix}$$

Por lo tanto, el sistema tiene solución si, y sólo si

$$4b_1 - b_2 + b_3 = 0$$

O sea

$$(x, y, z) \in Im\ T \Leftrightarrow 4x - y + z = 0$$

o bien,

$$Im\ T = \{(x, y, z) \mid 4x - y + z = 0\}$$

(subespacio de $\mathbf{R}^3$ que geométricamente representa un plano que pasa por el origen; el plano $4x - y + z = 0$).

Una base de este subespacio se puede obtener si se escribe el vector $(x, y, z) \in Im\, T$ como

$$(x, 4x + z, z) = x(1, 4, 0) + z(0, 1, 1)$$

los vectores $v_1 = (1, 4, 0)$ y $v_2 = (0, 1, 1)$ constituyen una base de $Im\, T$.

EJEMPLO 2 Considere ahora la transformación $T: M_{2\times 2} \to M_{3\times 1}$ dada por

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a + b + c + d \\ 2a + 3b + 4c + 6d \\ a + 3b + 5c + 9d \end{bmatrix}$$

Se deja al lector verificar que T es lineal.

Describa su núcleo y su imagen.

El vector $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}$ pertenece al núcleo de T si, y sólo si

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a + b + c + d \\ 2a + 3b + 4c + 6d \\ a + 3b + 5c + 9d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

de donde

$$a + b + c + d = 0$$

$$2a + 3b + 4c + 6d = 0$$

$$a + 3b + 5c + 9d = 0$$

Al resolver este sistema se obtienen los valores de a, b, c, d tales que $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in Ker\, T$.

Al proceder por el método de eliminación Gaussiana,

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 6 \\ 1 & 3 & 5 & 9 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -1 & -3 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

de donde

$$a = c + 3d$$

$$b = -2c - 4d$$

Por tanto,

$$Ker\, T = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2} \mid a = c + 3d, b = -2c - 4d\right\}$$

Una base de este subespacio se puede obtener escribiendo

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c + 3d & -2c - 4d \\ c & d \end{bmatrix} = c\begin{bmatrix} 1 & -2 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 3 & -4 \\ 0 & 1 \end{bmatrix}$$

Se verifica que los vectores $v_1 = \begin{bmatrix} 1 & -2 \\ 1 & 0 \end{bmatrix}$ y $v_2 = \begin{bmatrix} 3 & -4 \\ 0 & 1 \end{bmatrix}$ son linealmente independientes y por tanto (ya que ellos generan a $Ker\, T$) constituyen una base del núcleo de T.

Por otra parte el vector $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in M_{3\times 1}$ pertenece a la imagen de T si, y sólo si existe un vector $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}$ tal que

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a + b + c + d \\ 2a + 3b + 4c + 6d \\ a + 3b + 5c + 9d \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

es decir,

$$a + b + c + d = x_1$$

$$2a + 3b + 4c + 6d = x_2$$

$$a + 3b + 5c + 9d = x_3$$

Al resolver este sistema se obtiene

$$\begin{bmatrix} 1 & 1 & 1 & 1 & x_1 \\ 2 & 3 & 4 & 6 & x_2 \\ 1 & 3 & 5 & 9 & x_3 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -1 & -3 & 3x_1 - x_2 \\ 0 & 1 & 2 & 4 & x_2 - x_1 \\ 0 & 0 & 0 & 0 & 3x_1 - 2x_2 + x_3 \end{bmatrix}$$

Por lo tanto, el vector $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in M_{3\times 1}$ pertenece a $Im\, T$, si, y sólo si

$$3x_1 - 2x_2 + x_3 = 0$$

Es decir,

$$Im\, T = \left\{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in M_{3\times 1} \mid 3x_1 - 2x_2 + x_3 = 0\right\}$$

Una base de $Im\, T$ se puede obtener escribiendo

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 2x_2 - 3x_1 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

Se verifica que los vectores $v_1 = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$ y $v_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ son linealmente independientes y por lo tanto forman una base de *Im T*.

2.2. EL TEOREMA DE LA DIMENSIÓN

DEFINICIÓN 2.1 Sean V y U dos espacios vectoriales, V de dimensión finita y $T: V \to U$ una transformación lineal. Se define la *nulidad* de la transformación como la dimensión de su núcleo, y el *rango* de la transformación como la dimensión de su imagen.

EJEMPLO 3 Por ejemplo, para la transformación $T: \mathbf{R}^3 \to \mathbf{R}^3$

$$T(x, y, z) = (x + 2y + z, 3x + y - 2z, -x - 7y - 6z)$$

vista anteriormente se tiene

$$\text{Nulidad de } T = \dim Ker\ T = 1$$

$$\text{Rango de } T = \dim Im\ T = 2$$

y para la transformación $T: M_{2\times 2} \to M_{3\times 1}$

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a + b + c + d \\ 2a + 3b + 4c + 6d \\ a + 3b + 5c + 9d \end{bmatrix}$$

se tiene

$$\text{Nulidad de } T = \dim Ker\ T = 2$$

$$\text{Rango de } T = \dim Im\ T = 2$$

El siguiente teorema dice cómo están relacionadas la nulidad y el rango de una transformación con la dimensión de su dominio (en el caso de que éste sea un espacio vectorial de dimensión finita).

TEOREMA 2.2 (El teorema de la dimensión.)
Sean V y U dos espacios vectoriales, V de dimensión finita. Sea $T: V \to U$ una transformación lineal. Entonces

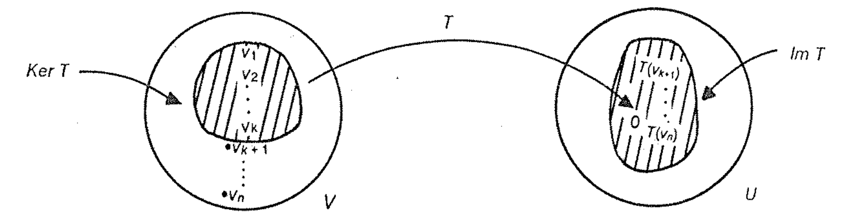
$$\text{rango de } T + \text{nulidad de } T = \dim V.$$

*DEMOSTRACIÓN** Sea $\beta_1 = \{v_1, v_2, \ldots, v_k\}$ una base de *Ker T*. Se sabe que existen vectores $v_{k+1}, \ldots, v_n$ en V tales que

$$\beta_2 = \{v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n\}$$

es una base de V (teorema 5.3 capítulo 3).

Se demostrará que $\beta_3 = \{T(v_{k+1}), \ldots, T(v_n)\}$ es una base de *Im T*. Esquemáticamente se tiene la siguiente situación:



Véase primeramente que los vectores de β_3 generan *Im T*. Sea $u \in Im\ T$. Existe entonces $v \in V$ tal que $T(v) = u$.

Para el vector $v \in V$ se escribe

$$v = \delta_1 v_1 + \delta_2 v_2 + \ldots + \delta_k v_k + \delta_{k+1} v_{k+1} + \ldots + \delta_n v_n$$

(pues β_2 es una base de V). Entonces

$$u = T(v) = T(\delta_1 v_1 + \delta_2 v_2 + \ldots + \delta_k v_k + \delta_{k+1} v_{k+1} + \ldots + \delta_n v_n)$$
$$= \delta_1 T(v_1) + \delta_2 T(v_2) + \ldots + \delta_k T(v_k) + \delta_{k+1} T(v_{k+1}) + \ldots + \delta_n T(v_n)$$

Pero $T(v_1) = \ldots = T(v_k) = 0$, pues, $v_1, \ldots, v_k$ son vectores de *Ker T*.

Entonces

$$u = \delta_{k+1} T(v_{k+1}) + \ldots + \delta_n T(v_n)$$

lo que prueba que $T(v_{k+1}), \ldots, T(v_n)$ generan a *Im T*.

Véase ahora que estos vectores son linealmente independientes. Tómese la combinación lineal

$$c_{k+1} T(v_{k+1}) + \ldots + c_n T(v_n) = 0$$

Se debe mostrar que $c_{k+1} = \ldots = c_n = 0$. La expresión anterior se puede reescribir como (usando el hecho de que T es lineal)

$$T(c_{k+1} v_{k+1} + \ldots + c_n v_n) = 0$$

lo cual dice entonces que

$$c_{k+1} v_{k+1} + \ldots + c_n v_n \in Ker\ T$$

*Se presenta la demostración en el caso en el que *Ker T* es un subespacio *no trivial* de V. La demostración en los casos en que $Ker\ T = \{0\}$ o $Ker\ T = V$ se deja como ejercicio para el lector.

Pero los vectores $v_1, v_2, \ldots, v_k$ constituyen una base de *Ker T*. Existen entonces escalares $c_1, c_2, \ldots, c_k$, tales que

$$c_{k+1} + v_{k+1} + \ldots + c_n v_n = c_1 v_1 + c_2 v_2 + \ldots + c_k v_k$$

expresión que se puede reescribir como

$$c_1 v_1 + c_2 v_2 + \ldots + c_k v_k + (-c_{k+1})v_{k+1} + \ldots + (-c_n)v_n = 0$$

Pero los vectores $v_1, v_2, \ldots, v_n$ constituyen una base V.

En particular ellos son linealmente independientes. Entonces la expresión anterior implica que

$$c_1 = c_2 = \ldots = c_k = c_{k+1} = \ldots = c_n = 0$$

Lo que muestra que $T(v_{k+1}), \ldots, T(v_n)$ son linealmente independientes y que por tanto constituyen una base de *Im T*.

Se tiene entonces que

$$\text{Nulidad de } T = \dim Ker\ T = k$$

$$\text{Rango de } T = \dim Im\ T = n - k$$

de modo que

Rango de T + nulidad de $T = n - k + k = n = \dim V$ como se quería demostrar.

Q.E.D.

EJEMPLO 4 Para las transformaciones lineales de los ejemplos dados anteriormente se tiene

$$T: \mathbf{R}^3 \to \mathbf{R}^3$$

$$T(x, y, z) = (x + 2y + z, 3x + y - 2z, -x - 7y - 6z)$$

$$\text{rango de } T + \text{nulidad } T = 2 + 1 = 3 = \dim \mathbf{R}^3$$

y

$$T: M_{2\times2} \to M_{3\times1}$$

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a + b + c + d \\ 2a + 3b + 4c + 6d \\ a + 3b + 5c + 9d \end{bmatrix}$$

$$\text{rango de } T + \text{nulidad de } T = 2 + 2 = 4 = \dim M_{2\times2}.$$

EJERCICIOS (SECCIÓN 2, CAPÍTULO 4)

1. Demuestre que el núcleo y la imagen de una transformación lineal son siempre conjuntos no vacíos.
2. Considere la transformación lineal $D: P \to P$ (P es el espacio vectorial de todos los polinomios) dada por $D(p) = p'$ (la derivada del polinomio p). Describa el núcleo de D.
3. Considere la transformación lineal $T: C([-\pi, \pi]) \to \mathbf{R}$ dada por

$$T(f) = \int_{-\pi}^{\pi} f(x)\,dx$$

Compruebe que $\cos x \in Ker\ T$. Describa en general el tipo de funciones del espacio $C([-\pi, \pi])$ que pertenecen al núcleo de T.
4. Sea $T: V \to U$ la transformación cero, describa el núcleo y la imagen de T.
5. Sea $T: V \to V$ el operador identidad en V, describa el núcleo y la imagen de T.
6. Sea $T: V \to \mathbf{R}$ una transformación lineal del espacio vectorial V de dimensión 1 a los reales, suponga que existe un vector no nulo $v \in V$ tal que $T(v) \neq 0$. Demuestre entonces que $T(x) \neq 0$ *para todo* vector no nulo $x \in V$.
7. Sea $T: \mathbf{R}^2 \to \mathbf{R}^2$ la transformación lineal tal que $T((1, 1)) = (0, 0)$ y $T((0, 1)) = (1, 1)$, demuestre que tanto el núcleo como la imagen de T son rectas en el plano xy que pasan por el origen. Encuentre las ecuaciones de estas rectas.
8. Sea $T: \mathbf{R}^2 \to \mathbf{R}^2$ una transformación lineal tal que $T((3, 2)) = (0, 0)$ y $T((1, 3)) = v \neq 0$, pruebe que el núcleo de T es una recta en el plano xy que pasa por el origen. Encuentre su ecuación.
9. Sea $T: \mathbf{R}^3 \to \mathbf{R}^2$ una transformación lineal tal que $T((1, 1, 0)) = T((0, 2, 1)) = (0, 0)$ y $T((-1, 2, 4)) = v \neq 0$, demuestre que el núcleo de T es un plano en $\mathbf{R}^3$ que pasa por el origen. Encuentre su ecuación.
10. Sea $T: \mathbf{R}^n \to \mathbf{R}^n$ la transformación lineal $T(X) = AX$, en donde A es una matriz $m \times n$, ¿qué relación guardan el núcleo de T y el espacio solución del sistema homogéneo de ecuaciones lineales $AX = 0$?
11. Sea $T: M_{2\times2} \to M_{2\times2}$ la transformación lineal

$$T(A) = AB - BA \text{ en donde } B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Describa el núcleo de T.
12. Sea $T: M_{n\times n} \to M_{n\times n}$ la transformación lineal $T(A) = A - A^t$.
    a) Demuestre que T es una transformación lineal.
    b) Describa el núcleo y la imagen de T.
13. Considere la transformación lineal $T: \mathbf{R}^4 \to \mathbf{R}^2$

$$T(x_1, x_2, x_3, x_4) = (x_1 - x_2 + 2x_3 + 3x_4, x_1 - 3x_2 + 4x_3 + 5x_4)$$

   a) Demuestre que $v_1 = (-1, 1, 1, 0)$ y $v_2 = (-2, 1, 0, 1)$ forman una base para el núcleo de T.
   b) Compruebe que la imagen de T es $\mathbf{R}^2$.

c) Sea $\beta = \{u_1, u_2\}$ una base cualquiera de $\mathbf{R}^2$, demuestre que existen vectores v_3 y v_4 en $\mathbf{R}^4$ tales que $T(v_3) = u_1$ y $T(v_4) = u_2$.

d) Pruebe que $\beta = \{v_1, v_2, v_3, v_4\}$ es una base de $\mathbf{R}^4$.

14. Sea $T: \mathbf{R}^2 \to \mathbf{R}^3$ una transformación lineal cuyo núcleo es una recta en $\mathbf{R}^2$ que pasa por el origen. Demuestre que la imagen de T es una recta en $\mathbf{R}^3$ que pasa por el origen.

15. Sea $T: \mathbf{R}^3 \to \mathbf{R}^3$ una transformación lineal cuya imagen es un plano en $\mathbf{R}^3$ que pasa por el origen. Demuestre que el núcleo de T es una recta en $\mathbf{R}^3$ que pasa por el origen.

16. Sea $T: V \to V$ un operador lineal en el espacio vectorial V de dimensión finita, suponga que $Im\ T$ es un subespacio *propio* de V. Demuestre que existe al menos un vector $v \in V$ no nulo tal que $T(v) = 0$.

17. Sea $T: \mathbf{R}^2 \to V$ una transformación lineal de $\mathbf{R}^2$ al espacio vectorial V, suponga que $Im\ T$ es un subespacio de V de dimensión 2. Compruebe que no existe $v \in V$ no nulo tal que $T(v) = 0$.

18. Para cada una de las transformaciones lineales siguientes, determine bases para su núcleo y su imagen, y verifique en cada caso que se satisface el teorema de la dimensión.

a) $T: \mathbf{R}^2 \to \mathbf{R}^2$, $T(x, y) = (x + y, x - y)$

b) $T: \mathbf{R}^3 \to \mathbf{R}^2$, $T(x, y, z) = (x + y - z, x - y - z)$

c) $T: \mathbf{R}^3 \to \mathbf{R}^2$, $T(x, y, z) = (x + y + z, 2x + 2y + 2z)$

d) $T: \mathbf{R}^2 \to \mathbf{R}^3$, $T(x, y) = (x + y, 5x + 5y, x - y)$

e) $T: \mathbf{R}^3 \to \mathbf{R}^3$, $T(x, y, z) = (x + y, x + z, y + z)$

f) $T: \mathbf{R}^4 \to \mathbf{R}^4$, $T(X) = AX$ en donde A es la matriz $\begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 5 & 8 & -2 & -7 \end{bmatrix}$

g) $T: \mathbf{R}^2 \to \mathbf{R}^4$, $T(X) = AX$ en donde A es la matriz $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}$

h) $T: P_3 \to P_2$, $T(a_0 + a_1x + a_2x^2 + a_3x^3) = a_0 + a_1 + a_2x + a_3x^2$

i) $T: P_2 \to P_3$, $T(p) = xp$

j) $T: P_2 \to P_4$, $T(p) = x^2p$

k) $T: M_{2\times2} \to M_{2\times2}$, $T(A) = A^t$

l) $T: M_{3\times3} \to \mathbf{R}$, $T(A) = tr\, A$

m) $T: M_{2\times2} \to M_{3\times2}$, $T(A) = BA$ en donde B es la matriz $\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 0 \end{bmatrix}$

n) $T: M_{2\times2} \to P_3$, $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + bx + cx^2 + dx^3$

3. REPRESENTACIÓN POR MEDIO DE MATRICES

Considérese la transformación $T: \mathbf{R}^n \to \mathbf{R}^m$ dada por

$$T(X) = AX \tag{3.1}$$

en donde A es la matriz de orden $m \times n$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{3.2}$$

y se está pensando en el vector $X = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$ como la matriz $n \times 1$.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

(similarmente la matriz $m \times 1$ AX es el vector en $\mathbf{R}^m$ imagen de X bajo T).

Se ha visto en la sección 1 que T es una transformación lineal. En este caso se dice que T es la transformación de "multiplicación por la matriz A".

Véase ahora que para las transformaciones lineales de esta forma resulta una tarea muy simple describir su núcleo y su imagen.

Primeramente observe que

$$Ker\ T = \{X \in \mathbf{R}^n \mid AX = 0\} = \text{espacio solución del sistema homogéneo } AX = 0$$

Es decir, para describir el núcleo de la transformación (3.1), se tienen que describir las soluciones del sistema homogéneo de m ecuaciones lineales con n incógnitas $AX = 0$.

También se afirma que

$$Im\ T = \text{espacio línea de } A^t$$

En efecto, supóngase que el vector $Y \in \mathbf{R}^m$ pertenece a $Im\ T$. Esto significa que existe $X \in \mathbf{R}^n$ tal que $T(X) = AX = Y$.

Más explícitamente, existen $x_1, x_2, \ldots, x_n$ (las coordenadas de X) tales que

$$A = \begin{bmatrix} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n \\ & & & \cdots & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n \end{bmatrix}$$

o bien,

$$Y = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \qquad (3.3)$$

Obsérvese que los vectores $(a_{1i}, a_{2i}, \ldots, a_{mi}) \in \mathbf{R}^m$ $i = 1, 2, \ldots, n$, son los vectores línea de A^t.

Resumiendo, la expresión (3.3) dice que el vector $Y \in \mathbf{R}^m$ pertenece a *Im T* si, y sólo si Y puede escribirse como combinación lineal de las vectores línea de A^t.

En otras palabras, Y pertenece a *Im T* si, y sólo si Y pertenece al espacio línea de A^t. Esto prueba la afirmación.

Entonces, se ha demostrado el siguiente teorema:

TEOREMA 3.1

Sea A una matriz $m \times n$ y sea $T: \mathbf{R}^n \to \mathbf{R}^m$ la transformación lineal $T(X) = AX$. Entonces

1) *Ker T* = espacio solución de $AX = 0$.
2) *Im T* = espacio línea de A^t.

EJEMPLO 1

Por ejemplo, sea $T: \mathbf{R}^3 \to \mathbf{R}^4$ dada por $T(X) = AX$ donde A es

$$A = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 3 \\ -3 & -14 & 8 \\ 3 & 4 & 2 \end{bmatrix}$$

Para describir *Ker T* resuelva el sistema homogéneo $AX = 0$.

$$\begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 3 \\ -3 & -14 & 8 \\ 3 & 4 & 2 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Las soluciones del sistema son entonces $x_1 = -2t$, $x_2 = t$, $x_3 = t$, $t \in \mathbf{R}$, de modo que

$$Ker\ T = \{(x_1, x_2, x_3) \mid x_1 = -2t, x_2 = t, x_3 = t, t \in \mathbf{R}\}$$

Una base para *Ker T* es $(-2, 1, 1)$ (tome $t = 1$), y entonces Nulidad de $T = 1$.

Describa ahora *Im T*. Para esto tome A^t y halle una base de su espacio línea

$$\begin{bmatrix} 1 & 2 & -3 & 3 \\ 3 & 1 & -14 & 4 \\ -1 & 3 & 8 & 2 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Entonces los vectores $Y_1 = (1, 0, -5, 1)$ y $Y_2 = (0, 1, 1, 1)$ constituyen una base de *Im T* (teorema 7.2 capítulo 3), y éntonces Rango de $T = 2$.

En la siguiente subsección se verá que cuando se tiene una transformación lineal de un espacio vectorial V de dimensión finita a otro espacio vectorial U también de dimensión finita, se podrá ver, por medio de bases concretas de V y de U, a la transformación T como si fuera una transformación del tipo (3.1).

3.1. LA MATRIZ ASOCIADA A UNA TRANSFORMACIÓN LINEAL

Sean V y U dos espacios vectoriales de dimensión finita, dígase que $\dim V = n$ y $\dim U = m$.

Tómese bases β_1 de V y β_2 de U

$$\beta_1 = \{v_1, v_2, \ldots, v_n\}$$

$$\beta_2 = \{u_1, u_2, \ldots, u_m\}$$

Sea $T: V \to U$ una transformación lineal entre estos espacios.

Para el vector $v \in V$ existen escalares $x_1, x_2, \ldots, x_n$ tales que

$$v = x_1v_1 + x_2v_2 + \ldots + x_nv_n$$

Es decir,

$$[v]_{\beta_1} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

La imagen de v bajo T es el vector

$$T(v) = T(x_1v_1 + x_2v_2 + \ldots + x_nv_n) = x_1\,T(v_1) + x_2\,T(v_2) + \ldots + x_n\,T(v_n) = \sum_{j=1}^{n} x_j\,T(v_j) \qquad (3.4)$$

Cada vector $T(v_j)$, $j = 1, 2, \ldots, n$ se encuentra en U (más concretamente, en *Im T*), de modo que existen escalares $a_{1j}, a_{2j}, \ldots, a_{mj}$ tales que

$$T(v_j) = a_{1j}u_1 + a_{2j}u_2 + \ldots + a_{mj}u_m = \sum_{i=1}^{n} a_{ij}u_i, \quad j = 1, 2, \ldots, n \qquad (3.5)$$

Es decir,

$$[T(v_j)]_{\beta_2} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \qquad j = 1, 2, \ldots, n$$

Al sustituir las n expresiones (3.5) en (3.4) se obtiene

$$T(v) = \sum_{j=1}^{n} x_j T(v_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}u_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}x_j \right) u_i \tag{3.6}$$

En vista de la unicidad de la expresión del vector $T(v) \in U$ como combinación lineal de los vectores $u_1, u_2, \ldots, u_m$ de la base β_2 de U, se concluye que

$$(T(v))_{\beta_2} = \left(\sum_{j=1}^{n} a_{1j}x_j, \sum_{j=1}^{n} a_{2j}x_j, \ldots, \sum_{j=1}^{n} a_{mj}x_j \right)$$

o bien,

$$[T(v)]_{\beta_2} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{bmatrix}$$

Considérese la matriz $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$. Obsérvese que

$$A[v]_{\beta_1} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{bmatrix} = [T(v)]_{\beta_2}$$

Es decir,

$$[T(v)]_{\beta_2} = A[v]_{\beta_1} \tag{3.7}$$

La matriz A es tal que en su j-ésima columna se encuentran los elementos de la matriz de coordenadas del vector $T(v_j)$ (la imagen del j-ésimo vector de la base β_1 de V) con respecto de la base β_2 de U.

Esquemáticamente

$$A = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ [T(v_1)]_{\beta_2} & [T(v_2)]_{\beta_2} & \ldots & [T(v_n)]_{\beta_2} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

Según (3.7), esta matriz A tiene la propiedad de que multiplicada por la matriz de coordenadas del vector $v \in V$ con respecto a la base β_1, da por resultado la matriz de coordenadas del vector $T(v) \in U$ con respecto a la base β_2 de U.

A la matriz A se le llama *matriz de la transformación T con respecto a las bases β_1 y β_2* (A es pues la matriz *asociada* a T).

Obsérvese que los elementos que constituyen esta matriz asociada a la transformación T *depende* de las bases β_1 y β_2 elegidas para V y U respectivamente.

Para remarcar esta dependencia, se escribirá:

$$A = [T]_{\beta_1\beta_2}$$

El análisis anterior queda entonces resumido a la fórmula

$$[T(v)]_{\beta_1} = [T]_{\beta_2\beta_1}[v]_{\beta_2} \tag{3.8}$$

En el caso particular en que se tenga un operador lineal $T: V \to V$ y que la base del dominio y codominio sea la misma (dígase β), se escribirá a la matriz A (de orden $n = dim\ V$) como $[T]_\beta$ y se dirá que ésta es la matriz del operador lineal T respecto de la base β.

Así entonces, cada transformación lineal $T: V \to U$, donde V y U son espacios de dimensión finita, tiene asociada una matriz de orden $dim\ U \times dim\ V$ (la cual, se insiste, depende de las bases elegidas para V y U) de tal manera que T puede ser vista como una transformación de multiplicación por la matriz $[T]_{\beta_1\beta_2}$ (la matriz asociada a T) en el sentido de la expresión (3.8).

En otras palabras, se ha establecido una *función* F (que depende de las bases de V y U) del conjunto de transformaciones lineales de V a U al conjunto de matrices $m \times n$.

$$F: \left\{ \begin{array}{c} \text{Transformaciones lineales} \\ T: V \to U \\ dim\ V = n, dim\ U = m \end{array} \right\} \to \left\{ \begin{array}{c} \text{Matrices } A \\ \text{de orden} \\ m \times n \end{array} \right\}$$

$$F(T) = [T]_{\beta_1\beta_2}$$

En la sección 5 se retomará esta discusión para investigar un poco más el carácter de esta función F. Ahora véase algunos ejemplos.

EJEMPLO 2 Sea T la transformación

$$T: M_{2\times 2} \to M_{3\times 1}$$

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 2a + 3b - c \\ a + 2b + c + 2d \\ a + b - 3c - 4d \end{bmatrix}$$

Se deja al lector verificar que T es una transformación lineal. Tome las bases canónicas de β_1 y β_2 de $M_{2\times 2}$ y $M_{3\times 1}$ respectivamente, es decir

$$\beta_1 = \left\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right\}$$

$$\beta_2 = \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$$

Obtenga la matriz $[T]_{\beta_1\beta_2}$.

Primeramente vea cuáles son las imágenes $T(v_j), j = 1, 2, 3, 4$ de los vectores de la base $\beta_1 = \{v_1, v_2, v_3, v_4\}$ de $M_{2\times 2}$.

$$T\left(\begin{bmatrix}1&0\\0&0\end{bmatrix}\right)=\begin{bmatrix}2\\1\\1\end{bmatrix} \qquad T\left(\begin{bmatrix}0&1\\0&0\end{bmatrix}\right)=\begin{bmatrix}3\\2\\1\end{bmatrix}$$

$$T\left(\begin{bmatrix}0&0\\1&0\end{bmatrix}\right)=\begin{bmatrix}-1\\1\\-3\end{bmatrix} \qquad T\left(\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)=\begin{bmatrix}0\\2\\-4\end{bmatrix}$$

Obsérvese que en este caso, como se está tomando la base canónica β_2 de $M_{3\times 1}$, se tiene, $T(v_j) = [T(v_j)]_{\beta_2} j = 1, 2, 3, 4$. Es decir,

$$[T(v_1)]_{\beta_2} = \begin{bmatrix}2\\1\\1\end{bmatrix},\ [T(v_2)]_{\beta_2} = \begin{bmatrix}3\\2\\1\end{bmatrix},\ [T(v_3)]_{\beta_2} = \begin{bmatrix}-1\\1\\-3\end{bmatrix},\ [T(v_4)]_{\beta_2} = \begin{bmatrix}0\\2\\-4\end{bmatrix}$$

de modo que

$$[T]_{\beta_1\beta_2} = \begin{bmatrix}2&3&-1&0\\1&2&1&2\\1&1&-3&-4\end{bmatrix}$$

Al usar esta matriz, se puede obtener la imagen de cualquier vector $v \in M_{2\times 2}$. Basta escribir la matriz de coordenadas $[v]_{\beta_1}$ y hacer el producto $[T]_{\beta_1\beta_2}[v]_{\beta_1}$. El resultado será la matriz de coordenadas del vector $T(v)$ respecto de la base β_2 (que en este caso coincide con $T(v)$). Por ejemplo:

$$\left(T\begin{bmatrix}2&3\\4&-8\end{bmatrix}\right)_{\beta_2} = [T]_{\beta_1\beta_2}\left(\begin{bmatrix}2&3\\4&-8\end{bmatrix}\right)_{\beta_1} = \begin{bmatrix}2&3&-1&0\\1&2&1&2\\1&1&-3&-4\end{bmatrix}\begin{bmatrix}2\\3\\4\\-8\end{bmatrix} = \begin{bmatrix}9\\-2\\25\end{bmatrix}$$

EJEMPLO 3 Considere ahora la transformación $T: P_4 \to P_4$ dada por $T(p) = p'$ (la derivada del polinomio p). En la sección 1 se vio que T es un operador lineal. Tome la base canónica de P_4, $\beta = \{1, x, x^2, x^3, x^4\}$. Obtenga $[T]_\beta$.

Se tiene

$$T(1) = (1)' = 0 \Rightarrow (T(1))_\beta = (0, 0, 0, 0, 0)$$
$$T(x) = (x)' = 1 \Rightarrow (T(x))_\beta = (1, 0, 0, 0, 0)$$
$$T(x^2) = (x^2)' = 2x \Rightarrow (T(x^2))_\beta = (0, 2, 0, 0, 0)$$
$$T(x^3) = (x^3)' = 3x^2 \Rightarrow (T(x^3))_\beta = (0, 0, 3, 0, 0)$$
$$T(x^4) = (x^4)' = 4x^3 \Rightarrow (T(x^4))_\beta = (0, 0, 0, 4, 0)$$

Entonces

$$[T]_\beta = \begin{bmatrix}0&1&0&0&0\\0&0&2&0&0\\0&0&0&3&0\\0&0&0&0&4\\0&0&0&0&0\end{bmatrix}$$

Ésta es entonces "una matriz que deriva polinomios de grado menor o igual a 4" en el sentido de que multiplicando esta matriz por la matriz de coordenadas del polinomio $p \in P_4$ respecto de la base β, se obtiene el vector de coordenadas de la derivada de p respecto de la base β. Por ejemplo, si $p = 5 + 8x - 10x^2 + 6x^3 - 7x^4$, se tiene

$$[p]_\beta = \begin{bmatrix}5\\8\\-10\\6\\-7\end{bmatrix}$$

y

$$[T(p)]_\beta = [p']_\beta = [T]_\beta [p]_\beta = \begin{bmatrix}0&1&0&0&0\\0&0&2&0&0\\0&0&0&3&0\\0&0&0&0&4\\0&0&0&0&0\end{bmatrix}\begin{bmatrix}5\\8\\-10\\6\\-7\end{bmatrix} = \begin{bmatrix}8\\-20\\18\\-28\\0\end{bmatrix}$$

O sea que

$$p' = 8(1) + (-20)x + 18x^2 + (-28)x^3 + (0)x^4$$
$$= 8 - 20x + 18x^2 - 28x^3$$

EJEMPLO 4 Como último ejemplo, considere la transformación $T: \mathbf{R}^3 \to \mathbf{R}^4$

$$T(x, y, z) = (x + 3y - z, 2x + y + 3z, -3x - 14y + 8z, 3x + 4y + 2z)$$

T es una transformación lineal. Más aún, si β_1 y β_2 son las bases canónicas de $\mathbf{R}^3$ y $\mathbf{R}^4$ respectivamente, la matriz de esta transformación respecto de β_1 y β_2 es

$$[T]_{\beta_1\beta_2} = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 3 \\ -3 & -14 & 8 \\ 3 & 4 & 2 \end{bmatrix}$$

de modo que T puede escribirse como

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 3 \\ -3 & -14 & 8 \\ 3 & 4 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Éste es precisamente el ejemplo que se estudia al inicio de esta sección.

Con esta misma transformación, obtenga $[T]_{\beta_1'\beta_2'}$ en donde

$$\beta_1' = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$$

$$\beta_2' = \{(0, 0, 1, 1), (0, 1, 0, 1), (1, 1, 0, 1), (1, 1, 1, 1)\}$$

Se verifica fácilmente (usando por ejemplo el teorema 4.3′ del capítulo 3), que β_1' y β_2' son bases de $\mathbf{R}^3$ y $\mathbf{R}^4$ respectivamente.

Escriba v_i ($i = 1, 2, 3$) para denotar a los vectores de β_1' y u_j ($j = 1, 2, 3, 4$) para denotar los vectores de β_2'.

Obtenga entonces $T(v_i)$, $i = 1, 2, 3$.

$$T(v_1) = T(1, 0, 0) = (1, 2, -3, 3)$$

$$T(v_2) = T(1, 1, 0) = (4, 3, -17, 7)$$

$$T(v_3) = T(1, 1, 1) = (3, 6, -9, 9)$$

Se tienen ahora que escribir estos vectores como combinación lineal de los vectores u_1, u_2, u_3, u_4.

Para ahorrar trabajo, escriba en general

$$(x_1, x_2, x_3, x_4) = c_1u_1 + c_2u_2 + c_3u_3 + c_4u_4$$

$$= c_1(0, 0, 1, 1) + c_2(0, 1, 0, 1) + c_3(1, 1, 0, 1) + c_4(1, 1, 1, 1)$$

Al hacer las operaciones indicadas obtenga el sistema

$$\begin{aligned} c_3 + c_4 &= x_1 \\ c_2 + c_3 + c_4 &= x_2 \\ c_1 + + c_4 &= x_3 \\ c_1 + c_2 + c_3 + c_4 &= x_4 \end{aligned}$$

Resolviéndolo con el método de eliminación Gaussiana queda

$$\left[\begin{array}{cccc|c} 0 & 0 & 1 & 1 & x_1 \\ 0 & 1 & 1 & 1 & x_2 \\ 1 & 0 & 0 & 1 & x_3 \\ 1 & 1 & 1 & 1 & x_4 \end{array}\right] \sim \cdots \sim \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & x_4 - x_2 \\ 0 & 1 & 0 & 0 & x_2 - x_1 \\ 0 & 0 & 1 & 0 & x_1 - x_2 - x_3 + x_4 \\ 0 & 0 & 0 & 1 & x_2 + x_3 - x_4 \end{array}\right]$$

de donde

$$[(x_1, x_2, x_3, x_4)]_{\beta_2'} = \begin{bmatrix} x_4 - x_2 \\ x_2 - x_1 \\ x_1 - x_2 - x_3 + x_4 \\ x_2 + x_3 - x_4 \end{bmatrix}$$

Entonces

$$[T(v_1)]_{\beta_2'} = [(1, 2, -3, 3)]_{\beta_2'} = \begin{bmatrix} 1 \\ 1 \\ 5 \\ -4 \end{bmatrix}, [T(v_2)]_{\beta_2'} = [(4, 3, -17, 7)]_{\beta_2'} = \begin{bmatrix} 4 \\ -1 \\ 25 \\ -21 \end{bmatrix}$$

$$[T(v_3)]_{\beta_2'} = [(3, 6, -9, 9)]_{\beta_2'} = \begin{bmatrix} 3 \\ 3 \\ 15 \\ -12 \end{bmatrix}$$

de modo que

$$[T]_{\beta_1'\beta_2'} = \begin{bmatrix} 1 & 4 & 3 \\ 1 & -1 & 3 \\ 5 & 25 & 15 \\ -4 & -21 & -12 \end{bmatrix}$$

Resulta natural preguntarse si las matrices

$$[T]_{\beta_1\beta_2} = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 3 \\ -3 & -14 & 8 \\ 3 & 4 & 2 \end{bmatrix} \text{ y } [T]_{\beta_1'\beta_2'} = \begin{bmatrix} 1 & 4 & 3 \\ 1 & -1 & 3 \\ 5 & 25 & 15 \\ -4 & -21 & -12 \end{bmatrix}$$

tienen alguna relación, pues, finalmente, ambas representan *la misma* transformación lineal $T: \mathbf{R}^3 \to \mathbf{R}^4$ dada por

$$T(x, y, z) = (x + 3y - z, 2x + y + 3z, -3x - 14y + 8z, 3x + 4y + 2z)$$

Se verá en la siguiente subsección que efectivamente estas matrices están íntimamente relacionadas por intermedio de matrices de cambios de base (véase subsección 6.1 capítulo 3).

Pero antes de pasar a este análisis, se desearía llamar la atención a otra pregunta que naturalmente se podría plantear en base al material desarrollado en esta sección:

Dada la transformación lineal $T: V \to U$, con V y U espacios vectoriales de dimensión finita, se ha visto que se puede asociar a ésta una matriz A por medio de bases concretas V y de U, de modo que la transformación T puede verse como una "multiplicación por la matriz A". Al principio de esta sección se vio que las transformaciones de este tipo (las que son multiplicaciones por matrices) tienen la ventaja de que su núcleo y su imagen se describen fácilmente en términos de la matriz correspondiente (teorema 3.1).

Esto invita entonces a investigar el núcleo y la imagen de la transformación $T: V \to U$ usando para ello "la" matriz que la representa. Sin embargo, es claro que este procedimiento está sumamente comprometido con la elección de las bases de V y U, pues ellas determinan la estructura de "la" matriz que representa a la transformación. Surge entonces la interrogante: ¿se logran descripciones equivalentes del núcleo y la imagen de una transformación $T: V \to U$ independientemente de la matriz que la represente, esto es, independientemente de las bases que se elijan para V y para U?

Se verá a continuación que la respuesta a esta pregunta es afirmativa.

Considérese entonces la transformación lineal $T: V \to U$, donde V es un espacio vectorial de dimensión n y U es un espacio de dimensión m.

Fíjese una base $\beta_1 = \{v_1, v_2, \ldots, v_n\}$ para V y una base $\beta_2 = \{u_1, u_2, \ldots, u_m\}$ para U.

Llámese A a la matriz $m \times n$ que representa a T respecto de la base β_1 y β_2 (esto es, $A = [T]_{\beta_1\beta_2}$).

Considérese también la transformación $T_A: \mathbf{R}^n \to \mathbf{R}^m$ dada por $T_A(X) = AX$.

Se quiere mostrar que tanto el núcleo como la imagen de T se puede describir trabajando con la transformación T_A (para la cual vale que $Ker\ T_A$ = espacio solución de $AX = 0$ e $Im\ T$ = espacio línea de A^T).

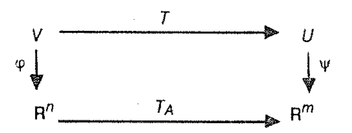
Considérense las funciones

$$\varphi: V \to \mathbf{R}^n$$
$$\psi: U \to \mathbf{R}^m$$

dadas por $\varphi(v) = (v)_{\beta_1}$ y $\psi(u) = (u)_{\beta_2}$. Es decir, φ y ψ asocian a los vectores de V y U, sus correspondientes vectores de coordenadas respecto de las bases β_1 y β_2, respectivamente.

En el teorema 6.4 del capítulo anterior (en su demostración) se vio que tanto φ como ψ son *isomorfismos* de espacios vectoriales, es decir, son funciones lineales biyectivas.

Se tiene entonces una situación como la siguiente



Lo que se quiere mostrar entonces es que

$$\varphi(Ker\, T) = Ker\, T_A$$
$$\psi(Im\, T) = Im\, T_A$$

Se expondrá detalladamente el argumento que conducirá a concluir la primera de estas afirmaciones

	$v \in Ker\, T$	
$\Leftrightarrow$	$T(v) = 0$	(definición de núcleo de T)
$\Leftrightarrow$	$\psi(T(v)) = \psi(0)$	($\Longrightarrow$ es obvio; $\Longleftarrow$ pues ψ es una biyección)
$\Leftrightarrow$	$(T(v))_{\beta_2} = 0$	(definición de ψ y hecho de que ψ es lineal)
$\Leftrightarrow$	$[T(v)]_{\beta_2} = 0$	(identificación del vector de coordenadas de $T(v)$ con una matriz de coordenadas)
$\Leftrightarrow$	$A[v]_{\beta_1} = 0$	(por la fórmula (3.8)).
$\Leftrightarrow$	$(v)_{\beta_1} \in Ker\, T_A$	(definición de núcleo de T_A e identificación de $[v]_{\beta_1}$ con $(v)_{\beta_1}$).
$\Leftrightarrow$	$\varphi(v) \in Ker\, T_A$	(definición de φ).

En resumen, se ha mostrado que $v \in Ker\, T \Leftrightarrow \varphi(v) \in Ker\, T_A$. Esto prueba entonces que $\varphi(Ker\, T) = Ker\, T_A$. La demostración de que $\psi(Im\, T) = Im\, T_A$ es similar y se deja como ejercicio para el lector.

EJEMPLO 5 Para ilustrar estos resultados, se retoma la transformación

$$T: M_{2\times 2} \to M_{3\times 1}$$

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a + b + c + d \\ 2a + 3b + 4c + 6d \\ a + 3b + 5c + 9d \end{bmatrix}$$

(véase subsección 2.1) para la cual ya se ha investigado "directamente" su núcleo y su imagen. Se vio que

$$Ker\ T = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2} \mid a = c + 3d,\ b = -2c - 4d \right\} \qquad (3.9)$$

$$Im\ T = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in M_{3\times 1} \mid 3x_1 - 2x_2 + x_3 = 0 \right\} \qquad (3.10)$$

Tome ahora las siguientes bases β_1 de $M_{2\times 2}$ y β_2 de $M_{3\times 1}$

$$\beta_1 = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \right\}$$

$$\beta_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

y describa *Ker T* e *Im T* por medio de la matriz $A = [T]_{\beta_1\beta_2}$.

Para construir la matriz A, halle la imagen bajo T de los vectores de la base β_1 y exprésela como combinación lineal de los vectores de la base β_2. Se tiene

$$T\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 5 \\ 4 \end{bmatrix} \qquad T\left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 12 \\ 15 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \\ -2 \end{bmatrix} \qquad T\left(\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Escriba en general

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_3\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Al hacer las operaciones indicadas se obtiene el sistema

$$\begin{aligned} c_1 + c_2 + c_3 &= x_1 \\ c_2 + c_3 &= x_2 \\ c_3 &= x_3 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} c_1 &= x_1 - x_2 \\ c_2 &= x_2 - x_3 \\ c_3 &= x_3 \end{aligned}$$

de modo que

$$\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right)_{\beta_2} = \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \\ x_3 \end{bmatrix}$$

y entonces

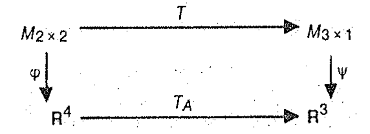
$$T\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right)_{\beta_2} = \begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix} \qquad T\left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\right)_{\beta_2} = \begin{bmatrix} -9 \\ -3 \\ 15 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right)_{\beta_2} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \qquad T\left(\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}\right)_{\beta_2} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$$

por lo que la matriz A es:

$$A = \begin{bmatrix} -3 & 1 & -9 & 0 \\ 1 & 1 & -3 & 2 \\ 4 & -2 & 15 & -1 \end{bmatrix}$$

Se tiene, pues, en este ejemplo, una situación como la que describe el siguiente diagrama:



en donde T_A: $\mathbf{R}^4 \to \mathbf{R}^3$ es $T_A(X) = AX$. Se describirá entonces el núcleo y la imagen de T (que corresponden a la parte "superior" del diagrama) trabajando con T_A (es decir, trabajando en la parte "inferior" del diagrama).

El núcleo de T_A es, como se sabe, el espacio solución del sistema homogéneo $AX = 0$. Resuelva entonces este sistema.

$$\begin{bmatrix} -3 & 1 & -9 & 0 \\ 1 & 1 & -3 & 2 \\ 4 & -2 & 15 & -1 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 3/2 & 1/2 \\ 0 & 1 & -9/2 & 3/2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

de donde

$$x_1 = -\frac{3}{2}x_3 - \frac{1}{2}x_4$$

$$x_2 = \frac{9}{2}x_3 - \frac{3}{2}x_4$$

Es decir,

$$Ker\ T_A = \left\{ (x_1, x_2, x_3, x_4) \in \mathbf{R}^4 \;\middle|\; x_1 = -\frac{3}{2}x_3 - \frac{1}{2}x_4,\ x_2 = \frac{9}{2}x_3 - \frac{3}{2}x_4 \right\}$$

Un vector representativo de $Ker\ T_A$ es entonces

$$\left(-\frac{3}{2}c-\frac{1}{2}d,\ \frac{9}{2}c-\frac{3}{2}d,\ c,\ d\right) \in \mathbf{R}^4,\ c, d \in \mathbf{R}$$

"Subiendo" en el diagrama, por medio de φ, se ve entonces que un vector representativo de $Ker\ T$ es:

$$\left(-\frac{3}{2}c-\frac{1}{2}d\right)\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}+\left(\frac{9}{2}c-\frac{3}{2}d\right)\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}+c\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}+d\begin{bmatrix}0 & 1\\ 1 & -1\end{bmatrix}$$

o bien,

$$\begin{bmatrix}-\frac{1}{2}(c+d) & 3c-d\\ -\frac{7}{2}c+\frac{5}{2}d & c-d\end{bmatrix}$$

Se comprueba fácilmente que esta descripción coincide exactamente con (3.9). La imagen T_A es el espacio línea de A^t.

$$\begin{bmatrix}-3 & 1 & 4\\ 1 & 1 & -2\\ -9 & -3 & 15\\ 0 & 2 & -1\end{bmatrix} \sim \cdots \sim \begin{bmatrix}1 & 0 & -3/2\\ 0 & 1 & -1/2\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$$

Una base de T_A está dada por los vectores $\left(1, 0, -\frac{3}{2}\right), \left(0, 1, -\frac{1}{2}\right)$.

Entonces un vector representativo de $Im\ T_A$ es del tipo

$$c_1\left(1, 0, -\frac{3}{2}\right)+c_2\left(0, 1, -\frac{1}{2}\right)=\left(c_1, c_2, -\frac{3}{2}c_1-\frac{1}{2}c_2\right)$$

De nuevo, "subiendo" en el diagrama, por medio de ψ, se ve que un vector representativo en $Im\ T$ es:

$$c_1\begin{bmatrix}1\\0\\0\end{bmatrix}+c_2\begin{bmatrix}1\\1\\0\end{bmatrix}+\left(-\frac{3}{2}c_1-\frac{1}{2}c_2\right)\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}-\frac{1}{2}c_1+\frac{1}{2}c_2\\ -\frac{3}{2}c_1+\frac{1}{2}c_2\\ -\frac{3}{2}c_1-\frac{1}{2}c_2\end{bmatrix}\quad c_1, c_2 \in \mathbf{R}$$

descripción que se comprueba con facilidad que coincide con (3.10).

Se sugiere al lector que rehaga este ejemplo tomando las bases canónicas de $M_{2\times 2}$ y $M_{3\times 1}$.

Pásese entonces ahora a estudiar cómo se afecta la matriz de una transformación lineal $T: V \to U$ cuando se cambian las bases de V y U.

3.2. CAMBIOS DE BASES

Sea $T: V \to U$ una transformación lineal, en donde V y U son espacios vectoriales de dimensión finita. Dígase que $n = dim\ V$ y $m = dim\ U$.

Sean β_1 y β_1' dos bases de V y β_2 y β_2' dos bases de U. Se quiere investigar la relación que existe entre las matrices

$$[T]_{\beta_1\beta_2} \quad \text{y} \quad [T]_{\beta_1'\beta_2'}$$

Para esto, recuérdese la subsección 6.1 del capítulo 3, que si X es un vector de V, existe una matriz inversible P de orden n (la matriz de cambio de base de β_1 a β_1') tal que

$$P[X]_{\beta_1} = [X]_{\beta_1'} \tag{3.11}$$

Similarmente, para cualquier vector dado Y en U, existe una matriz inversible Q de orden m (la matriz de cambio de base de β_2 a β_2') tal que

$$Q[Y]_{\beta_2} = [Y]_{\beta_2'} \tag{3.12}$$

Se sabe que para cualquier vector $v \in V$ se tiene

$$[T(v)]_{\beta_2} = [T]_{\beta_1\beta_2}[v]_{\beta_1} \tag{3.13}$$

y

$$[T(v)]_{\beta_2'} = [T]_{\beta_1'\beta_2'}[v]_{\beta_1'} \tag{3.14}$$

Según (3.12) se puede reescribir (el lado izquierdo de) la expresión (3.13) como

$$Q^{-1}[T(v)]_{\beta_2'} = [T]_{\beta_1\beta_2}[v]_{\beta_1}$$

Al sustituir (3.14) en esta última expresión se obtiene

$$Q^{-1}[T]_{\beta_1'\beta_2'}[v]_{\beta_1'} = [T]_{\beta_1\beta_2}[v]_{\beta_1}$$

Finalmente, según (3.11) se puede reescribir esta expresión como (sustituyendo $[v]_{\beta_1'}$ por $P[v]_{\beta_1}$).

$$Q^{-1}[T]_{\beta_1'\beta_2'}P[v]_{\beta_1} = [T]_{\beta_1\beta_2}[v]_{\beta_1}$$

Esta expresión es válida para todo vector $v \in V$. Entonces se puede concluir que*

$$[T]_{\beta_1\beta_2} = Q^{-1}[T]_{\beta_1'\beta_2'}I$$

Ésta es la relación que se quería establecer.

Se ha probado entonces el siguiente teorema:

TEOREMA 3.2

Sea $T: V \to U$ una transformación lineal del espacio vectorial V de dimensión n al espacio vectorial U de dimensión m. Sean β_1 y β_1' dos bases de V y β_2 y β_2' dos bases de U. Si P es la matriz de cambio de base de β_1 a β_1' (de orden n) y Q es la matriz de cambio de base de β_2 a β_2' (de orden m), entonces las matrices de la transformación T respecto de β_1 y β_2 y respecto de β_1' y β_2' están relacionadas por

$$[T]_{\beta_1\beta_2} = Q^{-1}[T]_{\beta_1'\beta_2'}P$$

o equivalentemente

$$[T]_{\beta_1'\beta_2'} = Q[T]_{\beta_1\beta_2}P^{-1} \qquad (3.15)$$

COROLARIO

Sea $T: V \to V$ un operador lineal y sean β y β' dos bases de V. Si P es la matriz de cambio de base de β y β' entonces

$$[T]_{\beta'} = P[T]_{\beta}P^{-1} \qquad (3.16)$$

DEMOSTRACIÓN En este caso $U = V$ y $\beta_1 = \beta_2$. El resultado se sigue entonces del teorema anterior.

Q.E.D.

EJEMPLO 6

Para ilustrar este teorema, tome la transformación del ejemplo dado en la página 306 de esta sección.

$$T: \mathbf{R}^3 \to \mathbf{R}^4$$

$$T(x, y, z) = (x + 3y - z, 2x + y + 3z, -3x - 14y + 8z, 3x + 4y + 2z)$$

Se tenía que

$$[T]_{\beta_1\beta_2} = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 3 \\ -3 & -14 & 8 \\ 3 & 4 & 2 \end{bmatrix}$$

*Si A y B son matrices $m \times n$ y X es una matriz $n \times 1$, se tiene: $AX = BX\ \forall X \in M_{n\times 1} \Rightarrow A = B$. Para ver esto, basta tomar $X = X_j, j = 1, 2, \ldots n$, donde X_j son las matrices de la base canónica de $M_{n\times 1}$ y ver que $AX_j = BX_j$ dice que las j-ésimas columnas de A y B coinciden.

de donde β_1 y β_2 eran las bases canónicas de $\mathbf{R}^3$ y $\mathbf{R}^4$ respectivamente. Es decir,

$$\beta_1 = \{(1, 0, 0,), (0, 1, 0), (0, 0, 1)\}$$

$$\beta_2 = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$$

Si β_1' y β_2' son las bases

$$\beta_1' = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$$

$$\beta_2' = \{(0, 0, 1, 1), (0, 1, 0, 1), (1, 1, 0, 1), (1, 1, 1, 1)\}$$

de $\mathbf{R}^3$ y $\mathbf{R}^4$ halle $[T]_{\beta_1'\beta_2'}$ usando el teorema 3.2.

Llame v_i, u_j, v_i', u_j', $i = 1, 2, 3, j = 1, 2, 3, 4$, a los vectores de las bases β_1, β_2, β_1' y β_2' respectivamente.

Si P es la matriz de cambio de base de β_1 a β_1' entonces P^{-1} es la matriz de cambio de base de β_1' a β_1. Esta última matriz es fácil de construir pues los vectores v_i' coinciden con sus vectores de coordenadas respecto de la base β_1 (pues β_1 es la base canónica de $\mathbf{R}^3$). Entonces

$$P^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Similarmente, construya la matriz Q^{-1} (matriz de cambio de base de β_2' a β_2) colocando en sus columnas las coordenadas de los vectores u_j' (que coinciden con sus vectores de coordenadas respecto de la base canónica β_2).

$$Q^{-1} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Obtenga $Q = (Q^{-1})^{-1}$

$$\left[\begin{array}{cccc|cccc} 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \sim \cdots \sim \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & -1 \end{array}\right]$$

de modo que

$$Q = \begin{bmatrix} 0 & -1 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix}$$

Entonces según la fórmula (3.15) se tiene que

$$[T]_{\beta_1'\beta_2'} = Q[T]_{\beta_1\beta_2}P^{-1} = \begin{bmatrix} 0 & -1 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 3 \\ -3 & -14 & 8 \\ 3 & 4 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 4 & 3 \\ 1 & -1 & 3 \\ 5 & 25 & 15 \\ -4 & -21 & -12 \end{bmatrix}$$

resultado que coincide con el que se había obtenido directamente en la página 307.

EJERCICIOS (SECCIÓN 3, CAPÍTULO 4)

1. Al usar el teorema 3.1, describa el núcleo y la imagen de las transformaciones lineales siguientes:

a) $T: \mathbf{R}^3 \to \mathbf{R}^2$ $T(X) = AX$ $A = \begin{bmatrix} 2 & 0 & -1 \\ -4 & 0 & 2 \end{bmatrix}$

b) $T: \mathbf{R}^4 \to \mathbf{R}^3$ $T(X) = AX$ $A = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 6 & 7 & 1 & 17 \\ 0 & 2 & 2 & 4 \end{bmatrix}$

c) $T: \mathbf{R}^4 \to \mathbf{R}^5$ $T(X) = AX$ $A = \begin{bmatrix} -1 & 1 & 3 & 2 \\ 2 & 4 & 14 & 10 \\ 2 & 1 & 4 & 3 \\ -10 & -4 & -10 & -10 \\ 3 & 1 & 1 & 2 \end{bmatrix}$

2. Encuentre la matriz asociada a cada una de las transformaciones lineales siguientes, respecto de las bases canónicas de los espacios correspondientes

a) $T: \mathbf{R}^2 \to \mathbf{R}^2$ $T(x, y) = (3x - 2y, 5x + y)$
b) $T: \mathbf{R}^3 \to \mathbf{R}^2$ $T(x, y, z) = (5x + 2y - z, 3x + 4y - 2z)$
c) $T: \mathbf{R}^3 \to \mathbf{R}^3$ $T(x, y, z) = (2x + y - z, 3x + 2y - 2z, 8x - y)$
d) $T: \mathbf{R}^2 \to \mathbf{R}^4$ $T(x, y) = (2x + y, x - 3y, x, y)$
e) $T: \mathbf{R}^4 \to \mathbf{R}^2$ $T(x, y, z, u) = (x + y, z + u)$
f) $T: \mathbf{R}^5 \to \mathbf{R}$ $T(x_1, x_2, x_3, x_4, x_5) = x_1 + x_2 + 2x_3 - 3x_4 + 8x_5$
g) $T: \mathbf{R}^6 \to \mathbf{R}^3$ $T(x_1, x_2, x_3, x_4, x_5, x_6) = (x_1 + x_4, x_2 + x_5, x_3 + x_6)$
h) $T: P_2 \to P_3$ $T(p) = xp$

i) $T: \mathbf{R}^2 \to P_1$ $T(a, b) = a + bx$
j) $T: P_3 \to \mathbf{R}^3$ $T(a_0 + a_1x + a_2x^2 + a_3x^3) = (2a_0 + a_1 + a_2, a_0 - 2a_2 + 3a_3, a_3)$
k) $T: P_1 \to P_4$ $T(p) = (x^3 + x^2 + x + 1)p$
l) $T: M_{2\times2} \to M_{2\times2}$ $T(A) = A^t$
m) $T: M_{2\times2} \to M_{3\times2}$ $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a+b & c \\ d & a+c \\ b+c & b+d \end{bmatrix}$
n) $T: M_{2\times2} \to \mathbf{R}$ $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + b + c + d$
o) $T: M_{3\times3} \to \mathbf{R}$ $T(A) = tr\, A$

3. Sean V y U espacios vectoriales de dimensión finita y sean β y β' bases de V y U respectivamente, ¿cuál es la matriz asociada a la transformación cero $0: V \to U$ respecto de las bases β y β'?
4. Sea V un espacio vectorial de dimensión finita y sea β una base de él, demuestre que la matriz asociada al operador identidad en V respecto de la base β es la matriz identidad.
5. Para cada una de las transformaciones lineales del ejercicio 2, describa su núcleo y su imagen por medio de alguna de sus matrices asociadas (teorema 3.1).
6. Considere la transformación lineal $T: M_{2\times2} \to M_{2\times2}$ dada por $T(A) = A - A^t$. En el ejercicio 12 de la sección anterior se demostró que el núcleo de T está constituido por las matrices simétricas (matrices tales que $A = A^t$) y que la imagen de T está constituida por las matrices antisimétricas (matrices tales que $B = -B^t$). Obtenga esta misma conclusión usando:
   a) la matriz de la transformación T respecto de la base canónica de $M_{2\times2}$.
   b) la matriz de la transformación T respecto de la base β de $M_{2\times2}$ dada por
   $$\beta_1 = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$$
7. Considere la transformación lineal $T: M_{2\times2} \to M_{2\times2}$ dada por $T(A) = BA$ en donde B es la matriz $\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$. Obtenga la matriz T respecto de la base canónica de $M_{2\times2}$. Describa el núcleo y la imagen de T.
8. Considere la transformación lineal $T: M_{2\times2} \to M_{2\times2}$ dada por $T(A) = BA$ en donde B es la matriz $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Sea C la matriz de T respecto de la base canónica de $M_{2\times2}$. Demuestre que $det\, C = (ad - bc)^2$. Concluya entonces que existe $A \in M_{2\times2}, A \neq 0$ tal que $T(A) = 0$ si, y sólo si B es una matriz no inversible.
9. Sea T la transformación lineal $T: M_{3\times3} \to M_{3\times3}$ dada por $T(A) = AB - BA$, en donde B es la matriz
   $$B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$
   Por medio de la matriz asociada a T respecto de la base canónica de $M_{3\times3}$ describa la estructura de las matrices que conmutan con B.

(*Sugerencia*: se trata de describir el núcleo de T).

② 10. Sean W_1 y W_2 dos subespacios del espacio vectorial V de dimensión finita tales que $V = W_1 \oplus W_2$. Sea $\beta_1 = \{v_1, v_2, \ldots, v_k\}$ una base de W_1 y $\beta_2 = \{u_1, u_2, \ldots, u_r\}$ una base de W_2.

a) Demuestre que $\beta = \{v_1, \ldots, v_k, u_1, \ldots, u_r\}$ es una base de V.

b) Sea $T: V \to V$ la transformación

$$T(c_1v_1 + \ldots + c_kv_k + d_1u_1 + \ldots + d_ru_r) = c_1v_1 + \ldots + c_kv_k$$

Compruebe que T es una transformación lineal.

c) Determine la estructura de la matriz de T respecto de la base β.

d) Describa el núcleo y la imagen de T.

② 11. Sea V un espacio vectorial de dimensión finita y sean V_1, V_2, V_3 y V_4 subespacios de V tales que $V = V_1 \oplus V_2 \oplus V_3 \oplus V_4$. Sea $T: V \to V$ la transformación tal que $T(v) = v$ si $v \in V_1 \cup V_3$ y $T(v) = kv$ si $v \in V_2 \cup V_4$.

a) Demuestre que T es una transformación lineal.

b) Si β_i es una base de V_i, $i = 1, 2, 3, 4$, demuestre que $\beta = \bigcup_{i=1}^{4} \beta_i$ es una base de V.

c) Describa la estructura de la matriz del operador T respecto de la base β.

12. Sea V el espacio vectorial $C^1(\mathbf{R})$. Considere el operador lineal $D: V \to V$ $D(f) = f'$ (la derivada de f). Obtenga la matriz de D respecto de la base dada de V (la constituida por los vectores que generan a V) en cada uno de los siguientes casos:

a) $V = \mathcal{L}(e^x, e^{-x})$

b) $V = \mathcal{L}(e^x, e^{2x}, e^{3x})$

c) $V = \mathcal{L}(e^x, e^{2x}, xe^x)$

d) $V = \mathcal{L}(e^x, xe^x, x^2e^x, x^3e^x)$

e) $V = \mathcal{L}(\operatorname{sen} x, \cos x)$

f) $V = \mathcal{L}(\operatorname{sen} x, \cos x, x \operatorname{sen} x, x \cos x)$

g) $V = \mathcal{L}(e^x \operatorname{sen} x, e^x \cos x)$

① 13. Sea $T: V \to U$ una transformación lineal del espacio vectorial V al espacio vectorial U. Sean $\beta_1 = \{v_1, v_2, \ldots, v_m\}$ y $\beta_2 = \{u_1, u_2, \ldots, u_n\}$ bases de V y U, respectivamente. Sea A la matriz de la transformación T respecto de las bases β_1 y β_2.

a) ¿Cómo se altera la matriz A si en β_1 se intercambian de posición los vectores v_i y v_j?

b) ¿Cómo se altera la matriz A si en β_2 se intercambian de posición los vectores u_i y u_j?

14. Considere el operador lineal $T: V \to V$ en donde V es un espacio vectorial de dimensión 4. Suponga que la matriz de T respecto de la base $\beta = \{v_1, v_2, v_3, v_4\}$ es

$$[T]_\beta = \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 2 & 1 & 1 \\ 3 & 2 & 1 & 4 \\ 0 & 2 & 4 & -1 \end{bmatrix}$$

Encuentre la matriz de T respecto de cada una de las siguientes bases de V:

a) $\beta_1 = \{v_2, v_1, v_4, v_3\}$

b) $\beta_2 = \{v_1, v_1 + v_2, v_1 + v_2 + v_3, v_1 + v_2 + v_3 + v_4\}$

c) $\beta_3 = \{v_1 + v_2, v_2 + v_3, v_3 + v_4, v_4\}$

15. Sea V un espacio vectorial de dimensión 3 y U un espacio vectorial de dimensión 4. Sean $\beta_1 = \{v_1, v_2, v_3\}$ y $\beta_2 = \{u_1, u_2, u_3, u_4\}$ bases de V y de U respectivamente. Sea $T: V \to U$ la transformación lineal tal que

$$T(v_1) = 2u_1 - 3u_2 + u_3 - u_4$$
$$T(v_2) = u_1 + u_2 + u_3 + u_4$$
$$T(v_3) = u_1 - 2u_3$$

a) determine la matriz de T respecto de las bases β_1 y β_2.

b) determine la matriz de T respecto de las bases

$$\beta_1' = \{v_1 + v_2 + v_3, 3v_1 - 2v_2 + 2v_3, v_3\}$$
$$\beta_2' = \{u_1 - u_2, u_1 + 2u_2 + 3u_4, u_3 - 2u_4, u_2 + u_3 + u_4\}$$

de V y U respectivamente.

16. Considere la transformación lineal $T: P_1 \to P_3$ dada por $T(p) = (x^2 + x + 1)p$

a) Determine la matriz de T respecto de las bases canónicas de P_1 y P_3.

b) Determine la matriz de T respecto de las bases

$$\beta_1' = \{1, x + 1\}$$
$$\beta_2' = \{2, x + 2, (x + 2)^2, (x + 2)^3\}$$

c) Verifique que se satisface el teorema 3.2.

17. Sea $T: \mathbf{R}^3 \to \mathbf{R}^3$ el operador lineal dado por

$$T(x, y, z) = (2x - y + z, x + z, 3y - 2z)$$

a) Determine la matriz de T respecto de la base canónica de $\mathbf{R}^3$.

b) Determine la matriz de T respecto de la base

$$\beta' = \{(1, 2, -1), (3, 0, 1), (0, -4, 0)\}$$

c) Verifique que se satisface el corolario del teorema 3.2.

18. La matriz de la transformación lineal $T: \mathbf{R}^3 \to \mathbf{R}^3$ respecto de la base $\beta = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ de $\mathbf{R}^3$ es:

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 3 & 2 & 2 \\ 4 & 3 & 2 \end{bmatrix}$$

Encuentre la matriz de T respecto de la base $\hat{\beta} = \{(2, 1, 1), (3, 2, -4), (2, 3, 3)\}$

4. SEMEJANZA

En toda esta sección se considerarán solamente *operadores lineales* $T: V \to V$ en donde V es un espacio vectorial de dimensión finita, así como las matrices (cuadradas) asociadas a ellos.

Considérese el conjunto $M_{n \times n}$ de todas las matrices cuadradas de orden n. En este conjunto se va a definir una relación entre sus elementos, llamada (relación de) semejanza. Esta relación ayudará a entender mejor la correspondencia que se estableció en la sección anterior de transformaciones lineales —matrices.

DEFINICIÓN 4.1 Sean A y B dos matrices cuadradas de orden n. Se dice que A es *semejante* a B, si existe una matriz P (de orden n) inversible tal que $A = P^{-1}BP$.

EJEMPLO 1 Por ejemplo, si $T: V \to V$ es un operador lineal en V y β y β' son dos bases distintas de V, la fórmula (3.16) dice que la matriz $[T]_\beta$ es semejante a la matriz $[T]_{\beta'}$ pues

$$[T]_\beta = P^{-1}[T]_{\beta'}P$$

en donde P resulta ser la matriz de cambio de base de β a β'. En realidad, ése es el "único" (tipo de) ejemplo que se puede dar de matrices semejantes. Es decir, dadas dos matrices A y B de orden n en donde A es semejante a B, existe un operador $T: V \to V$ en donde V es un espacio vectorial de dimensión n y bases β y β' de V tales que $A = [T]_{\beta'}$ y $B = [T]_\beta$.

Para ver la validez de esta afirmación, se demostrará primeramente el siguiente lema técnico:

LEMA Sea P una matriz inversible de orden n y sea V un espacio vectorial de dimensión n. Si β es una base de V, entonces existe una única base β' de V tal que P es la matriz de cambio de base de β' a β (esto es, tal que $[v]_\beta = P[v]_{\beta'}\ \forall v \in V$).

DEMOSTRACIÓN Denótese por p_{ij}, $i,j = 1, 2, \ldots, n$ a los elementos de la matriz P. Escríbase $\beta = \{v_1, v_2, \ldots, v_n\}$.

Defínase

$$u_j = p_{1j}v_1 + p_{2j}v_2 + \ldots + p_{nj}v_n = \sum_{i=1}^{n} p_{ij}v_i \quad j = 1, 2, \ldots, n \tag{4.1}$$

Si se logra mostrar que $\beta' = \{u_1, u_2, \ldots, u_n\}$ es una base de V, ésta será entonces la (única, por la manera como se definieron los vectores u_j) base de V tal que P es la matriz de cambio de base de β' a β (pues los elementos de la j-ésima columna de P son los correspondientes elementos de las matrices de coordenadas u_j respecto de la base β —ver subsección 6.1 capítulo 3).

Basta entonces mostrar que los vectores $u_1, u_2, \ldots, u_n$ son linealmente independientes (corolario del teorema 5.3 capítulo 3).

Escríbase la combinación lineal

$$c_1u_1 + c_2u_2 + \ldots + c_nu_n = 0 \tag{4.2}$$

Al sustituir las expresiones (4.1) en (4.2) queda

$$c_1\sum_{i=1}^{n} p_{i1}v_i + c_2\sum_{i=1}^{n} p_{i2}v_i + \ldots + c_n\sum_{i=1}^{n} p_{in}v_i = 0$$

o

$$\sum_{j=1}^{n} c_j \sum_{i=1}^{n} p_{ij}v_i = 0$$

o

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{n} p_{ij}c_j\right)v_i = 0 \tag{4.3}$$

Los vectores $v_1, v_2, \ldots, v_n$ forman por hipótesis, una base de V. En particular ellos son linealmente independientes. Entonces (4.3) implica que

$$\sum_{j=1}^{n} p_{ij}c_j = 0 \qquad i = 1, 2, \ldots, n$$

o bien, en términos matriciales, $PC = 0$, en donde

$$C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_3 \end{bmatrix}$$

Siendo P inversible, el sistema homogéneo $PC = 0$ tiene sólo la solución trivial. O sea que $c_1 = c_2 = \ldots = c_n = 0$, lo que prueba que $u_1, u_2, \ldots, u_n$ son linealmente independientes como se quería.

Q.E.D.

TEOREMA 4.1 Sean A y B dos matrices de orden n tales que A es semejante a B. Entonces existe un operador lineal $T: V \to V$, en donde V es un espacio vectorial de dimensión n, y bases β y β' de V tales que $A = [T]_{\beta'}$ y $B = [T]_\beta$.

DEMOSTRACIÓN Escríbase b_{ij}, $i,j=1, 2, \ldots, n$ para denotar a los elementos de la matriz B. Sea V un espacio vectorial de dimensión n y sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de él. Considérense los vectores

$$w_j = \sum_{i=1}^{n} b_{ij}v_i \qquad j = 1, 2, \ldots, n$$

Según el teorema 1.1 existe una (única) transformación lineal $T: V \to V$ tal que

$$T(v_j) = w_j = \sum_{i=1}^{n} b_{ij}v_i \qquad j = 1, 2, \ldots, n$$

Obsérvese que $[T]_\beta = B$.

Por otra parte, como A es semejante a B, existe una matriz Q inversible de orden n tal que $A = Q^{-1}BQ$. Según el lema anterior existe una base β' de V tal que Q es la matriz de cambio de base de β' a β. Por tanto, $P = Q^{-1}$ es la matriz de cambio de base de β a β'. Como

$$A = Q^{-1}BQ = P[T]_\beta P^{-1}$$

la fórmula (3.16) dice que $A = [T]_{\beta'}$.

Q.E.D.

En resumen, se tiene el siguiente corolario:

COROLARIO

Sean A y B dos matrices de orden n. A es semejante a B si, y sólo si A y B representan el mismo operador $T: V \to V$, en donde V es un espacio vectorial de dimensión n, respecto de dos bases β y β' de V.

TEOREMA 4.2

La relación de semejanza en el conjunto de matrices $M_{n\times n}$ es una relación de equivalencia.

DEMOSTRACIÓN

a) Es claro que toda matriz $A \in M_{n\times n}$ es semejante a sí misma, pues se puede escribir $A = I_n^{-1} A I_n$.

b) Supóngase ahora que A es semejante a B. Existe entonces una matriz inversible P tal que $A = P^{-1}BP$. Escríbase $Q = P^{-1}$. Entonces $B = Q^{-1}AQ$, lo que muestra que B es semejante a A.

c) Si A es semejante a B y B es semejante a C, existen matrices P_1 y P_2 inversibles tales que $A = P_1^{-1}BP_1$ y $B = P_2^{-1}CP_2$. Entonces $A = P_1^{-1}(P_2^{-1}CP_2)P_1 = (P_1^{-1}P_2^{-1})C(P_2P_1) = (P_1P_2)^{-1}C(P_1P_2)$. Es decir, existe $Q = P_1P_2$ inversible tal que $A = Q^{-1}CQ$. Entonces A es semejante a C.

Q.E.D.

Así pues, en particular se puede decir "las matrices A y B son semejantes" sin temor a confusión.

La relación de semejanza permite entonces asociar en clases de equivalencia a todas las matrices que representan al *mismo* operador $T: V \to V$ en diferentes bases de V. Éste es un hecho muy importante que se usará en la sección 6 en relación al estudio de la inversibilidad de operadores lineales.

Para terminar esta sección, se desearía hacer un comentario sobre la perspectiva que se abre ante los resultados que se acaban de obtener (corolario del teorema 4.1 y teorema 4.2). Dado un operador lineal $T: V \to V$, se ha visto que su núcleo y su imagen se pueden describir usando la matriz que representa a éste en alguna base de V. Más aún, en la próxima sección se verá que la matriz que representa a un operador es "un fiel reflejo" (en el sentido de isomorfismos) de él, de modo que toda la información sobre el operador está contenida en su representación matricial. Es natural preguntarse entonces si T tiene alguna representación matricial simple, pues finalmente con ella es con la que se puede trabajar en sustitución del operador mismo. Por ejemplo, uno se preguntaría si T tiene asociada una matriz diagonal. Para ponerlo con las palabras de la relación de semejanza que se estudia en esta sección, uno se pregunta si en la clase de equivalencia de las matrices que representan a T existe alguna matriz de forma especialmente simple, dígase diagonal. Ésta es una pregunta sumamente interesante, a la que se dará respuesta en el capítulo 6.

EJERCICIOS (SECCIÓN 4, CAPÍTULO 4)

1. En el espacio $\mathbf{R}^2$ considere la base $\beta = \{(2,3), (5, -1)\}$. Sea P la matriz

$$P = \begin{bmatrix} 1 & 4 \\ 8 & 7 \end{bmatrix}$$

Determine la base β' de $\mathbf{R}^2$ tal que P es la matriz de cambio de base de β' a β.

2. En el espacio $\mathbf{R}^3$ considere la base $\beta = \{(1, 0, 0), (0, 2, 4), (0, 1, 1)\}$. Sea P la matriz

$$P = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 4 & 5 \\ 0 & 2 & 0 \end{bmatrix}$$

Determine la base β' de $\mathbf{R}^2$ tal que P es la matriz de cambio de base de β a β' (no de β a β).

3. Demuestre que si las matrices A y B son semejantes entonces $\det A = \det B$.

4. Con las matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

construya un contraejemplo que muestre que la afirmación recíproca del ejercicio anterior es falsa.

① 5. Demuestre que si las matrices A y B son semejantes, entonces $tr\ A = tr\ B$. ¿Es válida la afirmación recíproca?

6. Sea A la matriz $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$, demuestre que las siguientes matrices no son semejantes a A.

a) $\begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix}$ c) $\begin{bmatrix} 0 & -1 \\ 5 & 7 \end{bmatrix}$ e) $\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 4 \\ 3 & 5 \end{bmatrix}$ d) $\begin{bmatrix} 5 & 1 \\ 1 & 1 \end{bmatrix}$ f) $\begin{bmatrix} 5 & 6 \\ 6 & 1 \end{bmatrix}$

7. Sea $T: \mathbf{R}^3 \to \mathbf{R}^3$ la transformación lineal

$$T(x, y, z) = (x + y, x + z, y + z)$$

Sean β_1, β_2 y β_3 las bases de $\mathbf{R}^3$ dadas por

$$\beta_1 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$
$$\beta_2 = \{(2, 1, 3), (3, 2, 2), (1, -1, 4)\}$$
$$\beta_3 = \{(3, -1, 0), (2, 0, 3), (0, 1, 1)\}$$

Sea

P la matriz de cambio de base de β_1 a β_2
Q la matriz de cambio de base de β_2 a β_3
R la matriz de cambio de base de β_3 a β_1

a) Obtenga la matriz de T respecto de la base β_1.
b) Obtenga la matriz de T respecto de la base β_2.
c) Obtenga la matriz de T respecto de la base β_3.
d) Verifique que

$$[T]_{\beta_1} = P^{-1}[T]_{\beta_2}P$$
$$[T]_{\beta_2} = Q^{-1}[T]_{\beta_3}Q$$
$$[T]_{\beta_3} = R^{-1}[T]_{\beta_1}R$$
$$R = (QP)^{-1}$$

8. Demuestre que las matrices

$$A = \begin{bmatrix} 8 & 4 \\ 3 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 7 & 1 \\ 17 & 3 \end{bmatrix}$$

son semejantes.

(*Sugerencia*: use la idea de la demostración del teorema 4.1.)

9. Sea A una matriz inversible de orden n, demuestre que la clase de equivalencia (bajo la relación de semejanza entre matrices) en la que se encuentra A está constituida sólo por matrices inversibles.

10. Describa la clase de equivalencia (bajo la relación de semejanza entre matrices) en la que se encuentra la matriz: a) cero, b) identidad.

11. Sean A y B dos matrices semejantes, demuestre que para cualquier natural n, las matrices A^n y B^n son también semejantes.

12. Sea A una matriz nilpotente (véase ejercicio 14 de la sección 3, capítulo 1), suponga que la matriz B es semejante a A. Demuestre que B es nilpotente.

13. Sea A una matriz involutoria (véase ejercicio 26 de la sección 3, capítulo 1), suponga que la matriz B es semejante a A. Demuestre que B es involutoria.

14. Sea A una matriz idempotente (véase ejercicio 13 de la sección 3, capítulo 1), suponga que la matriz B es semejante a A. Demuestre que B es idempotente.

15. Los tres ejercicios anteriores demuestran que las matrices semejantes a una matriz nilpotente, involutoria, idempotente, son matrices nilpotentes, involutorias o idempotentes, respectivamente. Esto *NO* significa que todas las matrices nilpotentes, involutorias o idempotentes sean semejantes entre sí. Proporcione ejemplos concretos que muestren esta afirmación.

⑤ 16. En este ejercicio se considerará un bonito e interesante problema de la **teoría de números**, conocido como "el problema de la **suma de Gauss**": se trata de calcular el valor de la suma

$$S = \sum_{j=0}^{n-1} \omega^{j^2}$$

en donde n es un número impar y w es una raíz n-ésima primitiva de la unidad, es decir, w es el número complejo

$$w = \cos\frac{2\pi}{n} + i \operatorname{sen}\frac{2\pi}{n}$$

Éste es un problema no trivial del cual se tienen reportadas en la bibliografía diversas maneras de resolverlo (algunas de ellas muy técnicas y complicadas). En los incisos siguientes se presenta el esqueleto de un argumento, un poco largo, pero muy simple, que resuelve el problema. La herramienta utilizada es principalmente tomada del álgebra lineal, de modo que el lector que haya seguido el material presentado en el texto hasta esta sección puede completar los detalles del mismo. Éste es el objetivo del ejercicio.

a) Primeramente se calculará el módulo del número completo S. Justifique los siguientes pasos

$$S\bar{S} = \left(\sum_{j=0}^{n-1} \omega^{j^2}\right)\left(\sum_{k=0}^{n-1} w^{-k^2}\right) = \sum_{j,k=0}^{n-1} w^{(j-k)(j+k)} = \sum_{u,v=0}^{n-1} w^{uv}$$

$$= \sum_{v=0}^{n-1} w^{(0)(v)} + \sum_{u=1}^{n-1}\left(\sum_{v=0}^{n-1} w^{uv}\right) = \sum_{v=0}^{n-1} (1) + \sum_{u=1}^{n-1} (0) = n$$

Concluya entonces que el módulo de S es $|S| = \sqrt{n}$.

b) Defina el conjunto G como $G = \{1, w, w^2, \ldots, w^{n-1}\}$. Verifique que G posee n elementos distintos. Los números complejos w^a y w^b (a, b enteros) se identificarán como siendo el mismo si $a \equiv b \bmod n$ (es decir, si existe k entero tal que $a = b + kn$). Justifique este hecho. Demuestre las siguientes propiedades del conjunto G:

b1) $g \in G$ si, y sólo si g es una raíz n-ésima de la unidad. (*Sugerencia*: la parte "sólo si" de la afirmación es muy simple de probar. Para ver la parte "si",

observe que siendo g una raíz n-ésima de la unidad, existe k entero tal que $g = w^k$. Aplique el algoritmo de la división en los enteros para concluir que $k = qn + j$, $0 \leq j \leq n-1$. Identifique entonces a g con w^j).

b2) Si $g_1, g_2 \in G$ entonces $g_1 g_2 \in G$ (es decir, G es un conjunto cerrado bajo el producto).

b3) Dado $g \in G$, existe $g^{-1} \in G$ tal que $(g)(g^{-1}) = 1$.

b4) La suma de todos los elementos de G es igual a cero. Es decir

$$1 + w + w^2 + \ldots + w^{n-1} = 0$$

c) Considere el conjunto $C(G)$ formado por todas las funciones complejas definidas en el conjunto G, es decir, funciones del tipo $f: G \to C$. En este conjunto defina las operaciones de suma y de producto por números complejos de la manera natural

$$f_1 + f_2: G \to C \qquad (f_1 + f_2)(g) = f_1(g) + f_2(g)$$
$$cf: G \to C \qquad (cf)(g) = cf(g) \quad (c \in C)$$

Si ahora usa la palabra "escalar" para designar a un número complejo, demuestre que el conjunto $C(G)$ con estas operaciones de suma y producto por escalares es un espacio vectorial (complejo).

d) Considere el conjunto $\beta_1 = \{f_0, f_1, \ldots, f_{n-1}\} \subset C(G)$ en donde

$$f_i(w^j) = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases} \qquad i, j = 0, 1, \ldots, n-1$$

Se demostrará que este conjunto es una base de $C(G)$. Tome primeramente la combinación lineal $c_0 f_0 + c_1 f_1 + \ldots + c_{n-1} f_{n-1} = 0$ (la función cero de $C(G)$). Evalúe esta expresión w^j $(0 \leq j \leq n-1)$ para concluir que $c_j = 0$. Esto prueba la independencia lineal β_1. Demuestre ahora que dada $f \in C(G)$, ésta puede escribirse como

$$f = \sum_{i=1}^{n-1} f(w^i) f_i$$

lo que prueba que el espacio generado por β_1 es $C(G)$. Concluya entonces que $C(G)$ es un espacio vectorial de dimensión n.

e) Sea s un número entero. Defina la función $e_s: G \to C$ como $e_s(w^r) = w^{sr}$. Considere el conjunto $\beta_2 = \{e_0, e_1, \ldots, e_{n-1}\} \subset C(G)$. Verifique las siguientes propiedades de los elementos de este conjunto

e1) $e_s(w^{r_1} w^{r_2}) = e_s(w^{r_1}) e_s(w^{r_2})$

e2) $|e_s(w^r)| = 1 \quad \forall s, r = 0, 1, \ldots, n-1$

e3) $\displaystyle\sum_{r=0}^{n-1} e_k(w^r) e_s(w^{-r}) = \begin{cases} 0 \text{ si } k \neq s \\ n \text{ si } k = s \end{cases}$

(*Sugerencia* para e3): use el resultado del inciso $b4$))

f) Demuestre que β_2 es un conjunto linealmente independiente $C(G)$. (*Sugerencia*: al evaluar la combinación lineal $c_0 e_0 + c_1 e_1 + \ldots + c_{n-1} e_{n-1} = 0$ en w^j $0 \leq j \leq n-1$, se obtiene $c_0 + c_1 w^j + \ldots + c_{n-1} w^{j(n-1)} = 0$. Se tiene así un sistema homogéneo de n ecuaciones lineales con n incógnitas $c_0, c_1, \ldots, c_{n-1}$. Verifique que el determinante de este sistema es un determinante de Vandermonde distinto de cero.) Concluya entonces que β_2 también es una base de $C(G)$.

g) Dada la función $f \in C(G)$, defina la función $\hat{f}: \beta_2 \to C$ como

$$\hat{f}(e_s) = \sum_{r=0}^{n-1} f(w^r) w^{-rs} = \sum_{r=0}^{n-1} f(w^r) e_s(w^{-r})$$

Al usar el hecho de que β_2 es una base de $C(G)$, escriba la expresión anterior como

$$\hat{f}(e_2) = \sum_{r=0}^{n-1} \left(\sum_{k=0}^{n-1} a_k e_k(w^r) \right) e_s(w^{-r})$$
$$= \sum_{k=0}^{n-1} a_k \left(\sum_{r=0}^{n-1} e_k(w^r) e_s(w^{-r}) \right)$$

en donde $a_0, a_1, \ldots, a_{n-1}$ son ciertos escalares que se evalúan como sigue: usando la expresión anterior y el resultado del inciso e3) se obtiene que

$$a_k = \frac{1}{n} \hat{f}(e_k)$$

Concluya entonces que la función $f \in C(G)$ se escribe en términos de la base β_2 como sigue

$$f = \frac{1}{n} \sum_{k=0}^{n-1} \hat{f}(e_k) e_k$$

h) Considere la función $T: C(G) \to C(G)$ dada por

$$T(f)(w^r) = \frac{1}{\sqrt{n}} \sum_{s=0}^{n-1} f(w^s) w^{rs} = \frac{1}{\sqrt{n}} \hat{f}(e_{-r})$$

Demuestre que T es un operador lineal en el espacio vectorial $C(G)$.

i) Dada la función $f \in C(G)$ defina la función $\check{f} \in C(G)$ como $\check{f}(w^r) = f(w^{-r})$. Demuestre que $T^2(f) = \check{f}$.

j) Defina los subconjuntos U y V del espacio vectorial $C(G)$ como

$$U = \{f \in C(G) \mid f = \check{f}\} \qquad V = \{f \in C(G) \mid f = -\check{f}\}$$

Demuestre que U y V son subespacios de $C(G)$.

k) Demuestre que $C(G) = U \oplus V$. (*Sugerencia*: dada $f \in C(G)$ considere las funciones $g, h \in C(G)$ definidas como $g = \frac{1}{2}(f + \check{f})$ y $h = \frac{1}{2}(f - \check{f})$. Verifique que $g \in U$, $h \in V$ y que $f = g + h$. Compruebe por último que $U \cap V = \{0\}$).

l) Considere las funciones $f_r \in C(G)$ (r un entero) definidas por

$$f_r(w^s) = \begin{cases} 1 \text{ si } r = s \text{ o } r = -s \\ 0 \text{ si } r \neq s \text{ y } r \neq -s \end{cases}$$

Verifique que $f_r \in U$ y que existen solamente $\frac{n+1}{2}$ funciones distintas de este tipo.

Demuestre que estas funciones constituyen una base de U. Concluya que

$$dim\ U = \frac{n+1}{2} \qquad dim\ V = \frac{n-1}{2}$$

m) Verifique que

m1) Si $f \in U$ entonces $T(f) \in U$ y $T^2(f) = f$.

m2) Si $f \in V$ entonces $T(f) \in V$ y $T^2(f) = -f$.

n) Considere los subconjuntos U_1, U_2 de U y V_1, V_2 de V definidos como

$$U_1 = \{f \in U \mid T(f) = f\} \qquad V_1 = \{f \in V \mid T(f) = if\}$$
$$U_2 = \{f \in U \mid T(f) = -f\} \qquad V_2 = \{f \in V \mid T(f) = -if\}$$

Demuestre que U_1 y U_2 son subespacios de U y que V_1 y V_2 son subespacios de V.

o) Demuestre que $U = U_1 \oplus U_2$ y que $V = V_1 \oplus V_2$. (*Sugerencia*: dada $f \in U$, escriba $f = \frac{1}{2}(f + T(f)) + \frac{1}{2}(f - T(f))$. Verifique que $\frac{1}{2}(f + T(f)) \in U_1$ y que $\frac{1}{2}(f - T(f)) \in U_2$. Dada $f \in V$ escriba $f = \frac{1}{2}(f - iT(f)) + \frac{1}{2}(f + iT(f))$. Verifique que $\frac{1}{2}(f - iT(f)) \in V_1$ y que $\frac{1}{2}(f + iT(f)) \in V_2$. Por último compruebe que $U_1 \cap U_2 = \{0\}$ y $V_1 \cap V_2 = \{0\}$.)

p) Sea $n_1 = dim\ U_1$, $n_2 = dim\ U_2$, $n_3 = dim\ V_1$, $n_4 = dim\ V_2$. Concluya de los incisos l) y o) que se tienen las fórmulas

$$n_1 + n_2 = \frac{n+1}{2} \qquad n_3 + n_4 = \frac{n-1}{2}$$

Si β_i es una base de U_i, $i = 1, 2$ y β_j es una base de V_j, $j = 1, 2$. Concluya que $\beta = \beta_1 \cup \beta_2 \cup \beta_3 \cup \beta_4$ es una base de $C(G)$.

q) Use el inciso n) y el ejercicio 11 de la sección 3 para concluir que la representación matricial del operador T: $C(G) \to C(G)$ definido en el inciso h) respecto de la base β de $C(G)$ del inciso anterior es

$$[T]_\beta = \begin{bmatrix} I_{n_1} & & & 0 \\ & -I_{n_2} & & \\ & & iI_{n_3} & \\ 0 & & & -iI_{n_4} \end{bmatrix}$$

en donde I_{n_j} es la matriz identidad de orden n_j, $j = 1, 2, 3, 4$.

r) Considere ahora la base $\beta_1 = \{f_0, f_1, \ldots, f_{n-1}\}$ de $C(G)$ establecida en el inciso d). Observando que

$$(Tf_k)(w^r) = \frac{1}{\sqrt{n}} \sum_{s=0}^{n-1} f_k(w^s) w^{rs} = \frac{1}{\sqrt{n}} w^{rk}$$

verifique que

$$T(f_k) = \frac{1}{\sqrt{n}} (f_0 + w^k f_1 + w^{2k} f_2 + \ldots + w^{(n-1)k} f_{n-1})$$

para $k = 0, 1, \ldots, n-1$. Use este hecho para concluir que la representación matricial del operador T respecto de esta base β_1 es

$$[T]_{\beta_1} = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & \ldots & 1 \\ 1 & w & \ldots & w^{n-1} \\ & & \ddots & \\ 1 & w^{n-1} & \ldots & w^{(n-1)^2} \end{bmatrix}$$

s) Obtenga las fórmulas para el determinante y la traza de las matrices que representan al operador T en las bases β y β_1

$$det\ [T]_\beta = (-1)^{n_2}(i)^{n_3 - n_4} \qquad tr\ [T]_\beta = n_1 - n_2 + i(n_3 - n_4)$$
$$det\ [T]_{\beta_1} = \frac{1}{n^{n/2}} \prod_{n-1 \geq j > k \geq 0} (w^j - w^k) \qquad tr\ [T]_{\beta_1} = \frac{1}{\sqrt{n}} S$$

en donde S es la suma de Gauss que se pretende calcular.

t) Use los ejercicios 3 y 5 de esta sección para establecer las igualdades

$$(-1)^{n_2} (i)^{n_3 - n_4} = \frac{1}{n^{n/2}} \prod_{n-1 \geq j > k \geq 0} (w^j - w^k) \qquad \text{(t1)}$$

$$\frac{1}{\sqrt{n}} S = n_1 - n_2 + i\,(n_3 - n_4) \qquad \text{(t2)}$$

(*Advertencia*: los resultados de esta sección y en general de todo el libro son válidos para espacios vectoriales reales. En este caso se están usando para un operador lineal en un espacio vectorial complejo. Acontece que prácticamente todos los teoremas establecidos hasta este momento se traducen sin cambios esenciales al caso de espacios vectoriales complejos. En particular, los resultados sobre la invariancia de la traza y el determinante para distintas representaciones matriciales de un operador lineal —o sea, el contenido de los ejercicios 5 y 3 respectivamente— se conservan sin cambio alguno en el caso en el que el operador lineal esté definido en un espacio complejo, situación que se presenta en este ejercicio).

u) Use la fórmula (t2) y el resultado del inciso a) para concluir que los números naturales n_1, n_2, n_3 y n_4 introducidos en el inciso p) son tales que solamente se puede tener una de las dos posibilidades siguientes:

u1) $n_1 = n_2$ y $|n_3 - n_4| = 1$ o u2) $n_3 = n_4$ y $|n_1 - n_2| = 1$

Más concretamente, demuestre que se tienen los dos casos siguientes:

Caso 1. Si $n \equiv 1 \bmod 4$, entonces $n_3 = n_4$ y $|n_1 - n_2| = 1$

Caso 2. Si $n \equiv 3 \bmod 4$, entonces $n_1 = n_2$ y $|n_3 - n_4| = 1$

(*Sugerencia*: si $n \equiv 1 \bmod 4$, entonces $n_1 + n_2 = \frac{1}{2}(n + 1) = 2k + 1$ para algún k entero, lo cual prueba que no se puede tener la opción $n_1 = n_2$. De manera análoga se ve que si $n \equiv 3 \bmod 4$, no se puede tener la opción $n_3 = n_4$.)

v) Demuestre que la fórmula (t1) se puede ver como

$$(-1)^{n_2} (i)^{n_3 - n_4} = (i)^{\frac{1}{2}n(n-1)}$$

(*Sugerencia*: $\displaystyle\prod_{n-1 \geq j > k \geq 0} (w^j - w^k) = \prod_{n-1 \geq j > k \geq 0} w^{\frac{1}{2}(j+k)} (w^{\frac{1}{2}(j-k)} - w^{-\frac{1}{2}(j-k)})$

$$= \prod_{n-1 \geq j > k \geq 0} w^{\frac{1}{2}(j+k)} \left(2i \operatorname{sen} (j-k) \frac{\pi}{n} \right) = (i)^{\frac{1}{2}n(n-1)} \prod_{n-1 \geq j > k \geq 0} w^{\frac{1}{2}(j+k)} 2 \operatorname{sen} (j-k) \frac{\pi}{n}$$

Observe que $\prod_{n-1 \geq j > k \geq 0} 2 \operatorname{sen} (j-k) \frac{\pi}{n} > 0$, pues el ángulo $(j-k)\frac{\pi}{n}$ con $n-1 \geq j > k \geq 0$ se encuentra siempre en el primero y segundo cuadrantes (siendo siempre

diferente de 0 y π). En particular, de la última igualdad se puede concluir que

$$(-1)^{n_2}(i)^{n_3-n_4} = (i)^{\frac{1}{2}n(n-1)} \prod_{n-1\geq j>k\geq 0} w^{\frac{1}{2}(j+k)}$$

Pruebe ahora que $\sum_{n-1\geq j>k\geq 0} (j+k) = \frac{1}{2}n(n-1)^2 = 2nk$, para algún k entero. Entonces se tiene

$$\prod_{n-1\geq j>k\geq 0} w^{\frac{1}{2}(j+k)} = w^{\frac{1}{2}(2nk)} = (w^n)^k = 1$$

probándose así la fórmula deseada).

w) Considere el caso 1 establecido en el inciso u), es decir; el caso en el que $n \equiv 1 \bmod 4$ y por lo tanto $n_3 = n_4$ y $|n_1 - n_2| = 1$. Demuestre que en este caso se tiene el hecho $n_3 = n_4$ y $n_1 - n_2 = 1$. (*Sugerencia*: considere los dos subcasos $n \equiv 1 \bmod 8$ y $n \equiv 5 \bmod 8$. Si $n \equiv 1 \bmod 8$ entonces $\frac{1}{2}n(n-1) = 4nk$ para algún k entero de modo que la fórmula establecida en el inciso anterior se concluye que $(-1)^{n_2} = (i)^{\frac{1}{2}n(n-1)} = 1$, lo que implica que n_2 debe ser par. Suponga que $n_1 - n_2 = -1$. Como $n_1 + n_2 = \frac{1}{2}(n + 1)$, concluya que esta suposición obliga a que $n_2 = 2k + 1$, contradiciendo el hecho de que n_2 es par. En este subcaso se debe tener entonces que $n_1 - n_2 = 1$. Un argumento completamente análogo muestra que también en el subcaso $n \equiv 5 \bmod 8$ se debe tener $n_1 - n_2 = 1$).

x) Use la fórmula (t2) y el resultado del inciso anterior para concluir que si $n \equiv 1 \bmod 4$, el valor de la suma de Gauss es $S = \sqrt{n}$.

y) Considere ahora el caso 2 establecido en el inciso u), es decir el caso en el que $n \equiv 3 \bmod 4$ y por lo tanto $n_1 = n_2$ y $|n_3 - n_4| = 1$. Demuestre que en este caso se tiene de hecho $n_1 = n_2$ y $n_3 - n_4 = 1$. (*Sugerencia*: en este caso se puede escribir la fórmula del inciso v) como

$$(i)^{\frac{1}{2}n(n-1)-2n_2-(n_3-n_4)} = 1$$

de donde $\frac{1}{2}n(n-1) - 2n_2 - (n_3 - n_4) = 4k$, para algún k entero. Como $n_1 = n_2$ y $n_1 + n_2 = \frac{1}{2}(n+1)$, se tiene $n_2 = \frac{1}{4}(n+1)$, y así

$$\begin{aligned} 4k &= \frac{1}{2}n(n-1) - 2n_2 - (n_3 - n_4) \\ &= \frac{1}{2}(4m+3)(4m+2) - (2)(\tfrac{1}{4})(4m+4) - (n_3 - n_4) \end{aligned}$$

en donde m es tal que $n = 4m + 3$, o sea que

$$(8m^2 + 8m + 1) - (n_3 - n_4) = 4k$$

Esta última expresión obliga a que $n_3 - n_4 = 1$, como se quería).

z) Use la fórmula (t2) y el resultado del inciso anterior para concluir que si $n \equiv 3 \bmod 4$ el valor de la suma de Gauss es $S = i\sqrt{n}$.

En resumen, la suma de Gauss es

$$S = \sum_{j=0}^{n-1} w^{j^2} = \begin{cases} \sqrt{n} & \text{si } n \equiv 1 \bmod 4 \\ i\sqrt{n} & \text{si } n \equiv 3 \bmod 4 \end{cases}$$

5. OPERACIONES CON TRANSFORMACIONES LINEALES

En esta sección se estudiará el conjunto de todas las transformaciones lineales de un espacio vectorial a otro. Se verán, por una parte, algunas operaciones entre sus elementos, y por otra, el reflejo de estas operaciones en el conjunto de matrices que representan a las transformaciones correspondientes.

5.1 SUMA Y PRODUCTO POR ESCALARES

Sean V y U dos espacios vectoriales (no necesariamente de dimensión finita) y sean

$$T_1: V \to U$$
$$T_2: V \to U$$

dos transformaciones lineales.

Defínase la suma de T_1 y T_2 y el producto de T_1 por el escalar $c \in \mathbb{R}$ de la manera usual

$$T_1 + T_2: V \to U$$
$$(T_1 + T_2)(v) = T_1(v) + T_2(v)$$

y

$$cT_1: V \to U$$
$$(cT_1)(v) = cT_1(v)$$

Siendo T_1 y T_2 transformaciones lineales, se afirma que $T_1 + T_2$ y cT_1 son también transformaciones lineales.

En efecto:

$$\begin{aligned} (T_1 + T_2)(v + v') &= T_1(v + v') + T_2(v + v') \\ &= T_1(v) + T_1(v') + T_2(v) + T_2(v') \\ &= T_1(v) + T_2(v) + T_1(v') + T_2(v') \\ &= (T_1 + T_2)(v) + (T_1 + T_2)(v') \end{aligned}$$

por lo que $T_1 + T_2$ preserva sumas de vectores en V. Si k es un escalar cualquiera

$$\begin{aligned} (T_1 + T_2)(kv) &= T_1(kv) + T_2(kv) \\ &= kT_1(v) + kT_2(v) = k(T_1(v) + T_2(v)) \\ &= k(T_1 + T_2)(v) \end{aligned}$$

por lo que $T_1 + T_2$ preserva también producto por escalares y entonces $T_1 + T_2$ es lineal.

Similarmente

$$(cT_1)(v + v') = cT_1(v + v') = c(T_1(v) + T_1(v'))$$
$$= cT_1(v) + cT_1(v')$$
$$= (cT_1)(v) + (cT_1)(v')$$

y si $k \in \mathbf{R}$

$$(cT_1)(kv) = cT_1(kv) = ckT_1(v)$$
$$= k(cT_1(v)) = k(cT_1)(v)$$

lo que muestra que cT_1 es lineal.

Denótese por $L(V, U)$ al conjunto de todas las transformaciones lineales de V a U. Entonces se han definido las operaciones

$$+: L(V, U) \times L(V, U) \to L(V, U)$$
$$\cdot: \quad \mathbf{R} \times L(V, U) \to L(V, U)$$

Es un sencillo ejercicio demostrar que con estas operaciones, el conjunto $L(V, U)$ es de hecho un *espacio vectorial*.

De aquí en adelante habrá que referirse entonces a $L(V, U)$ como el *espacio de transformaciones lineales de V a U.*

En el caso de que V y U son espacios de dimensión finita, se tienen algunos resultados sumamente interesantes acerca del espacio $L(V, U)$.

TEOREMA 5.1

Si V y U son espacios vectoriales de dimensión finita, entonces $L(V, U)$ es un espacio de dimensión finita y

$$\dim L(V, U) = (\dim V)(\dim U)$$

DEMOSTRACIÓN Sea $n = \dim V$ y $m = \dim U$. Tómese una base

$$\beta_1 = \{v_1, v_2, \ldots, v_n\} \text{ de } V \text{ y una base } \beta_2 = \{u_1, u_2, \ldots, u_m\}$$

de U.

Defínanse las funciones

$$f_{ij}: V \to U \qquad i = 1, 2, \ldots, m, j = 1, 2, \ldots, n$$

de la siguiente manera: para $v \in V$ escríbase

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n$$

Entonces

$$f_{ij}(v) = \lambda_j u_i$$

Es fácil verificar que cada una de estas mn funciones f_{ij} es de hecho una transformación lineal.

Se afirma que $\beta = \{f_{ij}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n\}$ constituye una base de $L(V, U)$ [llamada base canónica de $L(V, U)$].

Se verá primeramente que β genera a $L(V, U)$.

Tómese $T \in L(V, U)$ y véase que se puede expresar a T como una combinación lineal de los elementos de β.

Para cada $v_j \in \beta_1, j = 1, 2, \ldots, n$ escriba el vector $T(v_j)$, $\in U$ como combinación lineal de los vectores de la base β_2 de U.

$$T(v_j) = \alpha_{1j}u_1 + \alpha_{2j}u_2 + \ldots + \alpha_{mj}u_m, \qquad j = 1, 2, \ldots, n$$

Defínase la transformación $\tilde{T} \in \mathcal{L}(\beta)$ como

$$\tilde{T} = \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_{ij}f_{ij}$$

Se afirma que $\tilde{T} = T$.

Obsérvese primeramente que, por la manera como se definieron las transformaciones f_{ij} se tiene que

$$f_{ij}(v_k) = \begin{cases} u_i & \text{si } j = k \\ 0 & \text{si } j \neq k \end{cases}$$

Entonces

$$\tilde{T}(v_k) = \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_{ij}f_{ij}(v_k) = \sum_{i=1}^{m} \alpha_{ik}u_i = T(v_k)$$

Se tiene entonces que T y $\tilde{T}$ coinciden en todos los vectores de la base β_1 de V. Entonces $T = \tilde{T}$ (corolario del teorema 1.1) y por tanto $\mathcal{L}(\beta) = L(V, U)$.

Véase ahora que las transformaciones f_{ij} son linealmente independientes. Considérese la combinación lineal

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}f_{ij} = 0 \quad *$$

Es decir,

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}f_{ij}(v) = 0 \qquad \forall v \in V$$

*El símbolo 0 que aparece a la derecha de esta expresión se refiere al cero del espacio $L(V, U)$, el cual es la transformación cero $0: V \to U$, $0(v) = 0\ \forall v \in V$. (Véase sección 1.)

En particular para $v = v_k$, $k = 1, 2, \ldots, n$ se tiene

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}f_{ij}(v_k) = \sum_{i=1}^{m} c_{ik}u_i = 0 \qquad k = 1, 2, \ldots, n$$

Esta última expresión implica a su vez que $c_{ik} = 0$, $k = 1, 2, \ldots, n$, $i = 1, 2, \ldots, m$ (pues los vectores $u_1, u_2, \ldots, u_m$ son linealmente independientes). Por tanto, los vectores f_{ij} de β son linealmente independientes. Esto termina la demostración.

Q.E.D.

EJEMPLO 1

Por ejemplo, si $dim\ V = 2$ y $dim\ U = 3$, el espacio vectorial $L(V, U)$ tiene por base a las 6 transformaciones lineales f_{ij}: $V \to U$, $i = 1, 2, 3, j = 1, 2$ dadas por

$$f_{11}(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 u_1$$
$$f_{12}(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_2 u_1$$
$$f_{21}(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 u_2$$
$$f_{22}(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_2 u_2$$
$$f_{31}(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 u_3$$
$$f_{32}(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_2 u_3$$

en donde $\beta_1 = \{v_1, v_2\}$ y $\beta_2\ (u_1, u_2, u_3)$ son bases de V y de U, respectivamente. En este caso cualquier transformación $T \in L(V, U)$ puede expresarse como

$$T = \alpha_{11}f_{11} + \alpha_{12}f_{12} + \alpha_{21}f_{21} + \alpha_{22}f_{22} + \alpha_{31}f_{31} + \alpha_{32}f_{32}$$

en donde

$$T(v_1) = \alpha_{11}u_1 + \alpha_{21}u_2 + \alpha_{31}u_3$$
$$T(v_2) = \alpha_{12}u_1 + \alpha_{22}u_2 + \alpha_{32}u_3$$

Uno se podría preguntar por la forma de las matrices que representan a las transformaciones f_{ij}, $i=1, 2, 3, j=1, 2$ (del ejemplo anterior) con respecto de las bases β_1 y β_2.

Obténgase por ejemplo $[f_{11}]_{\beta_1\beta_2}$.

Se tiene

$$f_{11}(v_1) = u_1 \qquad \therefore \qquad [f_{11}(v_1)]_{\beta_2} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$f_{11}(v_2) = 0 \qquad \therefore \qquad [f_{11}(v_2)]_{\beta_2} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Entonces

$$[f_{11}]_{\beta_1\beta_2} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Unos cuantos cálculos más muestran que

$$[f_{12}]_{\beta_1\beta_2} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \ [f_{21}]_{\beta_1\beta_2} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \ [f_{22}]_{\beta_1\beta_2} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \ \text{etcétera.}$$

Es decir, que $[f_{ij}]_{\beta_1\beta_2}$ es la matriz de orden 3×2 que tiene por elementos solamente ceros, excepto el elemento de la i-ésima línea y j-ésima columna que es 1.

Se encontrará entonces, sorpresivamente, con las viejas conocidas matrices de la base canónica de $M_{3\times 2}$.

En el siguiente teorema se mostrará que la correspondencia que se estableció en la sección 3 entre transformaciones lineales y matrices

$$F: \left\{ \begin{array}{c} \text{Transformaciones lineales} \\ T: V \to U \\ dim\ V = n,\ dim\ U = m \end{array} \right\} \to \left\{ \begin{array}{c} \text{Matrices } A \\ \text{de orden} \\ m\times n \end{array} \right\} \qquad (5.1)$$

es de hecho un isomorfismo de espacios vectoriales (véase subsección 6.2 capítulo 3). Una de las propiedades de los isomorfismos, que ya se usaron en la sección 7 del capítulo 3 y que se mostrarán más adelante es que éstos llevan bases (de su dominio) en bases (de su codominio). Por lo tanto, el encuentro "sorpresivo" que se tuvo de las matrices en la base canónica de $M_{3\times 2}$ en el ejemplo anterior queda explicado por el siguiente teorema.

NOTA: El que el espacio $L(V, U)$ sea isomorfo a $M_{m\times n}$, es un hecho que se puede deducir de los teoremas 6.3 y 6.4 del capítulo 3 así como del teorema 5.1 de esta sección: se tiene que tanto $L(V, U)$ como $M_{m\times n}$ son isomorfos, a $\mathbf{R}^{mn}$, y por tanto, son isomorfos entre sí. El hecho que se presenta novedoso en este momento es que la función F definida en (5.1) sea un isomorfismo entre el espacio $L(V, U)$ y el espacio $M_{m\times n}$.

TEOREMA 5.2

Sea V un espacio vectorial de dimensión n y U un espacio vectorial de dimensión n con bases $\beta_1 = \{v_1, v_2, \ldots, v_n\}$ y $\beta_2\ \{u_1, u_2, \ldots, u_m\}$ respectivamente.

La función

$$F: L(V, U) \to M_{m\times n}$$

dada por

$$F(T) = [T]_{\beta_1\beta_2}$$

es un isomorfismo del espacio vectorial $L(V, U)$ al espacio vectorial $M_{m\times n}$.

DEMOSTRACIÓN Se debe mostrar que: (1) F es lineal, esto es, $F(T_1 + T_2) = F(T_1) + F(T_2)$, $T_1, T_2 \in L(V, U)$ y $F(cT) = cF(T)$, $c \in \mathbf{R}$, $T \in L(V, U)$; (2) F es una biyección.

Véase que F es lineal:

Los elementos de la j-ésima columna de la matriz $[T_1 + T_2]_{\beta_1\beta_2}$ son los elementos de la matriz [de coordenadas de $(T_1 + T_2)(v_j)$ respecto de la base β_2] $[(T_1 + T_2)(v_j)]_{\beta_2}$.

Pero

$$[(T_1 + T_2)(v_j)]_{\beta_2} = [T_1(v_j) + T_2(v_j)]_{\beta_2}$$
$$= [T_1(v_j)]_{\beta_2} + [T_2(v_j)]_{\beta_2}$$

Como la relación anterior es válida para toda columna j, se concluye que

$$F(T_1 + T_2) = [T_1 + T_2]_{\beta_1\beta_2} = [T_1]_{\beta_1\beta_2} + [T_2]_{\beta_1\beta_2} = F(T_1) + F(T_2)$$

Similarmente se prueba que $F(cT) = cF(T)$.

Véase que F es una biyección.

Sea $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ una matriz en $M_{m\times n}$. Defínase

$$w_j = a_{1j}u_1 + a_{2j}u_2 + \ldots + a_{mj}u_m \in U \qquad j = 1, 2, \ldots, n$$

El teorema 1.1 dice que:

a) *existe* una transformación T: $V \rightarrow U$ tal que

$$T(v_j) = w_j = a_{1j}u_1 + a_{2j}u_2 + \ldots + a_{mj}u_m \qquad j = 1, 2, \ldots, n$$

Se verifica fácilmente que $F(T) = [T]_{\beta_1\beta_2} = A$. Esto muestra que F es sobreyectiva.

b) esta transformación T es única.

Esto muestra que F es inyectiva.

Por lo tanto, f es un *isomorfismo*.

Q.E.D.

EJEMPLO 2 Por ejemplo, considérense las transformaciones T_1, T_2, $\mathbf{R}^2 \rightarrow \mathbf{R}^2$ dadas por

$$T_1(x, y) = (3x + 2y, 5x - y)$$
$$T_2(x, y) = (-4x + 5y, -x + 7y)$$

La suma de T_1 y T_2 es la transformación $T_1 + T_2$: $\mathbf{R}^2 \rightarrow \mathbf{R}^2$

$$(T_1 + T_2)(x, y) = T_1(x, y) + T_2(x, y)$$
$$= (3x + 2y, 5x - y) + (-4x + 5y, -x + 7y)$$
$$= (-x + 7y, 4x + 6y)$$

Si se toma la base de $\mathbf{R}^2$ $\beta = \{(1, 0), (1, 1)\}$, se tiene

$$[T_1]_\beta = \begin{bmatrix} -2 & 1 \\ 5 & 4 \end{bmatrix} \qquad [T_2]_\beta = \begin{bmatrix} -3 & -5 \\ -1 & 6 \end{bmatrix}$$

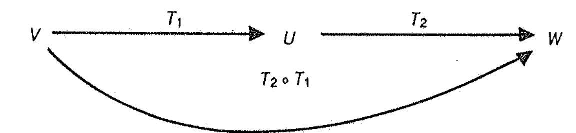
y por tanto

$$[T_1 + T_2]_\beta = [T_1]_\beta + [T_2]_\beta = \begin{bmatrix} -2 & 1 \\ 5 & 4 \end{bmatrix} + \begin{bmatrix} -3 & -5 \\ -1 & 6 \end{bmatrix} = \begin{bmatrix} -5 & -4 \\ 4 & 10 \end{bmatrix}$$

como puede verificarse directamente.

5.2 COMPOSICIÓN

Sean V, U y W espacios vectoriales, y sean T_1: $V \rightarrow U$ y T_2: $U \rightarrow W$ transformaciones lineales. Se puede formar la composición $T_2 \circ T_1$: $V \rightarrow W$ definida por $(T_2 \circ T_1)(v) = T_2(T_1(v))$, $v \in V$.



Se afirma que $T_2 \circ T_1$: $V \rightarrow W$ es una transformación lineal.

En efecto

$$(T_2 \circ T_1)(v + v') = T_2(T_1(v + v')) = T_2(T_1(v) + T_1(v'))$$
$$= T_2(T_1(v)) + T_2(T_1(v')) = (T_2 \circ T_1)(v) + (T_2 \circ T_1)(v')$$

y si $c \in \mathbf{R}$

$$(T_2T_1)(cv) = T_2(T_1(cv)) = T_2(cT_1(v)) = cT_2(T_1(v))$$
$$= c(T_2 \circ T_1)(v)$$

En el siguiente teorema se establecen las propiedades más importantes de la composición de transformaciones lineales:

TEOREMA 5.3

1) Si $T_1 \in L(V, U)$, $T_2 \in L(U, W)$, $T_3 \in L(W, R)$, entonces

$$T_3 \circ (T_2 \circ T_1) = (T_3 \circ T_2) \circ T_1 \qquad (\in L(V, R))$$

2) Si $T_1 \in L(V, U)$ y $T_2, T_3 \in L(U, W)$, entonces

$$(T_2 + T_3) \circ T_1 = T_2 \circ T_1 + T_3 \circ T_1 \qquad (\in L(V, W))$$

3) Si $T_1 \in L(U, W)$ y $T_2, T_3 \in L(V, U)$ entonces

$$T_1 \circ (T_2 + T_3) = T_1 \circ T_2 + T_1 \circ T_3 \qquad (\in L(V, W))$$

4) Si $c \in \mathbf{R}$ y $T_1 \in L(V, U)$ y $T_2, \in L(U, W)$, entonces

$$c(T_2 \circ T_1) = (cT_2) \circ T_1 = T_2 \circ (cT_1) \qquad (\in L(V, W))$$

DEMOSTRACIÓN Se trata de un ejercicio de simple aplicación de la definición. Se hará la demostración de 1) y 2) y se dejarán como ejercicio para el lector los ejemplos 3) y 4).

Sea $v \in V$, entonces

$$\begin{aligned}(T_3 \circ (T_2 \circ T_1))(v) &= T_3((T_2 \circ T_1(v))) = T_3(T_2(T_1(v)))\\ &= (T_3 \circ T_2)(T_1(v)) = ((T_3 \circ T_2) \circ T_1)(v)\end{aligned}$$

de donde se sigue 1).

Pruébese 2).

$$\begin{aligned}((T_2 + T_3) \circ T_1)(v) &= (T_2 + T_3)(T_1(v)) = T_2(T_1(v)) + T_3(T_1(v))\\ &= (T_2 \circ T_1)(v) + (T_3 \circ T_1)(v)\\ &= ((T_2 \circ T_1) + (T_3 \circ T_1))(v)\end{aligned}$$

y por tanto

$$(T_2 + T_3) \circ T_1 = T_2 \circ T_1 + T_3 \circ T_1$$

Q.E.D.

Supóngase ahora que V, U y W son espacios vectoriales de dimensión finita, dígase $dim\ V = n$, $dim\ U = r$, $dim\ W = m$. Tómense bases $\beta_1 = \{v_1, v_2, \ldots, v_n\}$, $\beta_{2'} = \{u_1, u_2, \ldots, u_r\}$ y $\beta_3 = \{w_1, w_2, \ldots, w_m\}$ de V, U y W respectivamente, y considérense las transformaciones lineales T_1 y T_2.

$$V \xrightarrow{\ T_1\ } U \xrightarrow{\ T_2\ } W$$

Tanto T_1 como T_2 y $T_2 \circ T_1$ tienen matrices asociadas respecto de las bases correspondientes: $[T_1]_{\beta_1\beta_2}$ es una matriz $dim\ U \times dim\ V = r \times n$, $[T_2]_{\beta_2\beta_3}$ es una matriz $dim\ W \times dim\ U = m \times r$ y $[T_2 \circ T_1]_{\beta_1\beta_3}$ es una matriz $dim\ W \times dim\ V = m \times n$.

Escríbase

$$[T_1]_{\beta_1\beta_2} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ & \cdot & \cdot & \cdot & \\ a_{r1} & a_{r2} & \ldots & a_{rn}\end{bmatrix} \qquad [T_2]_{\beta_2\beta_3} = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1r}\\ b_{21} & b_{22} & \ldots & b_{2r}\\ & \cdot & \cdot & \cdot & \\ b_{m1} & b_{m2} & \ldots & b_{mr}\end{bmatrix}$$

$$[T_2 \circ T_1]_{\beta_1\beta_3} = \begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1n}\\ c_{12} & c_{22} & \ldots & c_{2n}\\ & \cdot & \cdot & \cdot & \\ c_{m1} & c_{m2} & \ldots & c_{mn}\end{bmatrix}$$

Entonces

$$(T_2 \circ T_1)(v_j) = c_{1j}w_1 + c_{2j}w_2 + \ldots + c_{mj}w_m = \sum_{i=1}^{n} c_{ij}w_i$$

Pero, por otra parte

$$\begin{aligned}(T_2 \circ T_1)(v_j) &= T_2(T_1(v_j)) = T_2(a_{1j}u_1 + a_{2j}u_2 + \ldots + a_{rj}u_r)\\ &= T_2\left(\sum_{k=1}^{r} a_{kj}u_k\right)\\ &= \sum_{k=1}^{r} a_{kj}T_2(u_k)\\ &= \sum_{k=1}^{r} a_{kj}(b_{1k}w_1 + b_{2k}w_2 + \ldots + b_{mk}w_m)\\ &= \sum_{k=1}^{r} a_{kj} \sum_{i=1}^{m} b_{ik}w_i = \sum_{i=1}^{m}\left(\sum_{k=1}^{r} b_{ik}a_{kj}\right) w_i\end{aligned}$$

En vista de la unicidad de la representación del vector $(T_2 \circ T_1)(v_j)$ como combinación lineal de los vectores $w_1, w_2, \ldots, w_m$ se concluye que

$$c_{ij} = \sum_{k=1}^{r} b_{ik}a_{kj}$$

lo cual dice entonces que la matriz asociada a la composición $T_2 \circ T_1$ (respecto de las bases β_1 y β_3) no es más que el producto de las matrices asociadas a cada una de las transformaciones T_2 y T_1* (respecto de las bases correspondientes). O sea

$$[T_2 \circ T_1]_{\beta_1\beta_3} = [T_2]_{\beta_2\beta_3}\,[T_1]_{\beta_1\beta_2} \qquad (5.2)$$

EJEMPLO 3 Por ejemplo, sean $T_1: \mathbf{R}^3 \to \mathbf{R}^2$ y $T_2: \mathbf{R}^2 \to \mathbf{R}^2$, las **transformaciones** dadas por

$$T_1(x, y, z) = (2x + 6y - 3z, 5x - 3y + 8z)$$
$$T_2(x, y) = (2x + 5y, -x + 6y)$$

Entonces $T_2 \circ T_1: \mathbf{R}^3 \to \mathbf{R}^2$ está dada por

$$\begin{aligned}(T_2 \circ T_1)(x, y, z) &= T_2(T_1(x, y, z)) = T_2(2x + 6y - 3z, 5x - 3y + 8z)\\ &= (2(2x + 6y - 3z) + 5(5x - 3y + 8z), -(2x + 6y - 3z) + 6(5x - 3y + 8z))\\ &= (29x - 3y + 34\,z, 28x - 24y + 51z)\end{aligned}$$

*Véase entonces que la definición de producto de matrices que se dio en el capítulo 1 corresponde, según este elegante resultado, a la representación matricial de la composición de dos **transformaciones lineales**.

Tómese la base de $\mathbf{R}^3$ $\beta_1 = \{(1, 1, 1), (1, 0, 1), (0, 1, 1)\}$ y en $\mathbf{R}^2$ tómese la base **canónica** $\beta_2 = \{(1, 0), (0, 1)\}$

Entonces

$$T_1(1, 1, 1) = (5, 10), \quad T_1(1, 0, 1) = (-1, 13), \quad T_1(0, 1, 1) = (3, 5)$$

de modo que

$$[T_1]_{\beta_1\beta_2} = \begin{bmatrix} 5 & -1 & 3 \\ 10 & 13 & 5 \end{bmatrix}$$

Similarmente

$$T_2(1, 0) = (2, -1) \qquad \text{y} \qquad T_2(0, 1) = (5, 6)$$

de modo que

$$[T_2]_{\beta_2} = \begin{bmatrix} 2 & 5 \\ -1 & 6 \end{bmatrix}$$

Por lo tanto, según la fórmula (5.2) (con $\beta_3 = \beta_2$)

$$[T_2 \circ T_1]_{\beta_1\beta_2} = \begin{bmatrix} 2 & 5 \\ -1 & 6 \end{bmatrix}\begin{bmatrix} 5 & -1 & 3 \\ 10 & 13 & 5 \end{bmatrix} = \begin{bmatrix} 60 & 63 & 31 \\ 55 & 79 & 27 \end{bmatrix}$$

lo cual se comprueba directamente observando que de la fórmula que define a $T_2 \circ T_1$ se obtiene

$$(T_2 \circ T_1)(1, 1, 1) = (60, 55)$$
$$(T_2 \circ T_1)(1, 0, 1) = (63, 79)$$
$$(T_2 \circ T_1)(0, 1, 1) = (31, 27)$$

La fórmula (5.2) toma un aspecto más simple cuando T_1 y T_2 son operadores lineales en el espacio vectorial V. En el siguiente teorema se resume el análisis anterior para este caso particular.

TEOREMA 5.4

Sea V un espacio vectorial de dimensión n, y sean $T_1, T_2 \in L(V, V)$. Si β es una base de V se tiene

$$[T_1 \circ T_2]_\beta = [T_1]_\beta \, [T_2]_\beta$$

APÉNDICE. ÁLGEBRAS REALES

En este apéndice se van a reconsiderar algunos de los resultados que se obtuvieron en esta sección, bajo la perspectiva de una nueva estructura algebraica más general que la de espacio vectorial: la estructura de álgebra.

DEFINICIÓN

Un *álgebra* sobre el campo de los reales (o un álgebra real) es un espacio vectorial real A en el cual se tiene definida una operación adicional, llamada *multiplicación de vectores*, la cual asocia a cada par de vectores $v_1, v_2 \in A$ un nuevo vector $v_1v_2 \in A$ llamado *producto de* v_1 *y* v_2 de modo que se satisfacen las siguientes propiedades:

1) La multiplicación es asociativa

$$v_1(v_2v_3) = (v_1v_2)v_3 \qquad v_1, v_2, v_3 \in A$$

2) La multiplicación es distributiva respecto de la suma

$$v_1(v_2 + v_3) = v_1v_2 + v_1v_3$$
$$(v_1 + v_2)v_3 = v_1v_3 + v_2v_3 \qquad v_1, v_2, v_3 \in A$$

3) Para $c \in \mathbf{R}$ se tiene

$$c(v_1v_2) = (cv_1)v_2 = v_1(cv_2) \qquad v_1, v_2 \in A$$

Si además existe un elemento $e \in A$ tal que $ev = ve = v \; \forall \, v \in A$, se dice que A es un *álgebra con unidad*, y al elemento $e \in V$ se le llama *elemento unidad* (o identidad) de A. Por otra parte, si la multiplicación en el álgebra A cumple la condición adicional $v_1v_2 = v_2v_1 \; \forall \, v_1, v_2 \in A$, se dice que A es un *álgebra conmutativa*.

La estructura de álgebra es entonces más rica que la estructura de espacio vectorial, pues un álgebra A además de ser espacio vectorial tiene la operación de multiplicación entre sus elementos.

A continuación se verán dos ejemplos de espacios vectoriales (con los que ya se ha trabajado anteriormente) que de hecho son álgebras.

El espacio vectorial $M_{n \times n}$ de las matrices cuadradas de orden n, con la multiplicación de matrices definida en el capítulo 1 es un álgebra con unidad (la matriz I_n es el elemento unidad de $M_{n \times n}$).

En efecto, en el teorema 3.2 del capítulo 1 se prueba que para cualesquiera matrices cuadradas A, B, C del mismo orden y cualquier escalar $c \in \mathbf{R}$ se tiene

$$A(BC) = (AB)C$$
$$A(B + C) = AB + AC$$
$$(A + B)C = AC + BC$$
$$c(AB) = (cA)B = A(cB)$$

Lo que muestra, según la definición anterior, que $M_{n \times n}$ es un es un álgebra. Obsérvese que esta álgebra no es conmutativa.

El espacio vectorial $L(V, V)$ de los operadores lineales en un espacio vectorial V es un álgebra, si se define en él la multiplicación de operadores en $L(V, V)$ como su composición. Es decir, el producto de $T_1, T_2 \in L(V, V)$ es (por definición) $T_1 \circ T_2$.

En efecto, del teorema 5.3 se puede concluir que para $T_1, T_2, T_3 \in L(V, V)$ y $c \in \mathbf{R}$

$$(T_1 \circ T_2) \circ T_3 = T_1 \circ (T_2 \circ T_3)$$
$$T_1 \circ (T_2 + T_3) = T_1 \circ T_2 + T_1 \circ T_3$$
$$(T_1 + T_2) \circ T_3 = T_1 \circ T_3 + T_2 \circ T_3$$
$$c(T_1 \circ T_2) = (cT_1) \circ T_2 = T_1 \circ (cT_2)$$

lo que muestra que $L(V, V)$ es un álgebra.

DEFINICIÓN Sean A_1 y A_2 dos álgebras reales. Se dice que la función $F: A_1 \to A_2$ es un *isomorfismo de álgebras* si F es una biyección y además, para todo $v_1, v_2 \in A$ se cumple

1) $F(v_1 + v_2) = F(v_1) + F(v_2)$
2) $F(cv_1) = cF(v_1),\ c \in \mathbf{R}$
3) $F(v_1v_2) = F(v_1)F(v_2)$

Es decir, la biyección F es un isomorfismo del álgebra A_1 al álgebra A_2, si F "respeta" las operaciones que se tienen definidas en el álgebra (manda sumas, producto por escalares y multiplicación de vectores en A_1 a sumas, producto por escalares y multiplicación de vectores en A_2). Obsérvese que las condiciones 1) y 2) de la definición anterior dicen (junto con el hecho de que F es biyectiva) que F es en particular un isomorfismo de espacios vectoriales.

El resultado principal de este apéndice, que no es más que una reformulación condensada de varias de las ideas desarrolladas en esta sección, se encuentran en el siguiente teorema:

TEOREMA Sea V un espacio vectorial de dimensión n y sea β una base de V. La función

$$F: L(V, V) \to M_{m \times n}$$
$$F(T) = [T]_\beta$$

es un isomorfismos del álgebra $L(V, V)$ al álgebra $M_{n \times n}$.

DEMOSTRACIÓN En el teorema 5.2 se probó que F es un isomorfismo de espacios vectoriales. Sólo falta verificar que F respeta las operaciones de multiplicación de vectores.

Es decir:

$$F(T_1 \circ T_2) = F(T_1)F(T_2)$$

Pero esta afirmación es precisamente el contenido del teorema 5.4.

Q.E.D.

EJERCICIOS (SECCIÓN 5, CAPÍTULO 4)

1. Sean T_1 y T_2 las transformaciones lineales de $\mathbf{R}^3$ a $\mathbf{R}^3$ dadas por

$$T_1(x, y, z) = (2x + y + z, x - y + z, 3x + 2y - 2z)$$
$$T_2(x, y, z) = (5x + z, 2y + 3z, x - y - z)$$

a) Evalúe $(T_1 + T_2)(2, 1, 3)$, $(3T_1)(1, 1, 1)$ y $(-2T_2)(2, 4, 0)$.
b) Encuentre una fórmula general para las transformaciones $T_1 + T_2$, cT_1 y cT_2 ($c \in \mathbf{R}$).

2. Demuestre que $L(V, U)$ es un espacio vectorial. ¿Cuál es el cero de este espacio?, ¿cuál es el inverso aditivo de una transformación lineal T en este espacio?

3. Determine explícitamente bases para los siguientes espacios vectoriales:

a) $L(\mathbf{R}^2, \mathbf{R})$
b) $L(\mathbf{R}^2, \mathbf{R}^2)$
c) $L(\mathbf{R}^2, \mathbf{R}^3)$
d) $L(\mathbf{R}^3, \mathbf{R}^2)$
e) $L(P_1, P_1)$
f) $L(P_1, P_2)$
g) $L(M_{2\times 2}, M_{2\times 2})$
h) $L(M_{1\times 2}, M_{2\times 1})$

4. Suponga que la transformación lineal $T: \mathbf{R}^2 \to \mathbf{R}^3$ tiene asociada respecto de las bases canónicas de $\mathbf{R}^2$ y $\mathbf{R}^3$ la siguiente matriz:

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 1 \\ 0 & 3 \end{bmatrix}$$

mientras que la transformación $T_2: \mathbf{R}^2 \to \mathbf{R}^3$ tiene asociada la matriz

$$B = \begin{bmatrix} -1 & 3 \\ 2 & 8 \\ -4 & 2 \end{bmatrix}$$

Determine las matrices asociadas a las transformaciones $S: \mathbf{R}^2 \to \mathbf{R}^3$ (respecto de la base canónica) si

a) $S = T_1 + T_2$
b) $S = 2T_1$
c) $S = T_1 + 3T_2$
d) $S = 5T_1 - 2T_2$

5. Suponga que la transformación $T_1: \mathbf{R}^2 \to \mathbf{R}^2$ tiene asociada la matriz

$$A = \begin{bmatrix} -1 & 4 \\ 2 & 3 \end{bmatrix}$$

respecto de la base $\beta = \{(2, 1), (3, -2)\}$ de $\mathbf{R}^2$, mientras que la transformación $T_2 : \mathbf{R}^2 \to \mathbf{R}^2$ tiene asociada la matriz

$$B = \begin{bmatrix} 5 & 3 \\ 1 & -1 \end{bmatrix}$$

respecto de la base $\beta' = \{(3, 1), (3, -2)\}$ de $\mathbf{R}^2$. Determine la matriz de la transformación $T_1 + T_2: \mathbf{R}^2 \to \mathbf{R}^2$ respecto de la base

a) β
b) β'
c) $\beta_1 = \{(0, 1), (3, -2)\}$
d) $\beta_2 = \{(2, 1), (-1, 1)\}$
e) $\beta_3 = \{(4, 1), (5, 2)\}$
f) $\beta_4 = \{(1, 0), (0, 1)\}$

6. Demuestre que la operación de composición de transformaciones lineales es una operación no conmutativa.

7. Sean T_1, T_2 y T_3 las transformaciones lineales de $\mathbf{R}^2$ a $\mathbf{R}^2$ dadas por

$$T_1(x, y) = (2x - y, x + 3y)$$
$$T_2(x, y) = (5x + 2y, 3x + 2y)$$
$$T_3(x, y) = (-x, -x, -y)$$

Determine expresiones explícitas para cada una de las siguientes transformaciones:

a) $T_1 \circ T_2$
b) $T_2 \circ T_1$
c) $T_1 \circ T_2 \circ T_3$
d) $T_2 \circ (3T_1) \circ T_3$
e) $T_2 \circ (T_1 + T_3)$
f) $T_1 \circ (T_1 + 2T_2) \circ (3T_2 - T_3)$

8. Compruebe que si $T_1 \in L(U, W)$ y $T_2, T_3 \in L(V, U)$, se tiene la siguiente fórmula para la transformación $T_1 \circ (T_1 + T_3) \in L(U, W)$

$$T_1 \circ (T_2 + T_3) = T_1 \circ T_2 + T_1 \circ T_3$$

(distributividad de la composición de transformaciones lineales respecto de su suma).

9. Demuestre que si $T_1 \in L(V, U)$ y $T_2 \in L(U, W)$ se tiene

$$c(T_2 \circ T_1) = (cT_2) \circ T_1 = T_2 \circ (cT_1) \qquad c \in \mathbf{R}$$

10. Dado el operador lineal $T: V \to V$ en el espacio vectorial V, se define el operador lineal $T^n: V \to V$ ($n \in \mathbf{R}$) como

$$T^n = T \circ T \circ \ldots \circ T \qquad (n \text{ veces})$$

Sea $T: \mathbf{R}^2 \to \mathbf{R}^2$ el operador lineal dado por $T(x, y) = (2x - y, 5x + 2y)$. Pruebe que el operador $T^2 - 4T + 9\,Id: \mathbf{R}^2 \to \mathbf{R}^2$ (Id es el operador identidad en $\mathbf{R}^2$) es el operador cero en $\mathbf{R}^2$.

11. Sea $T: \mathbf{R}^3 \to \mathbf{R}^3$ el operador lineal

$$T(x, y, z) = (x, 2y, 3z)$$

Demuestre que $T^3 - 6T^2 + 11T - 6Id$ es el operador cero de $\mathbf{R}^3$.

12. Sea $D: P_n \to P_n$ el operador derivación $D(p) = p'$, compruebe que $D^{n+1} = 0$ (el operador cero en P_n).

13. Sea $T: M_{n\times n} \to M_{n\times n}$ el operador lineal dado por $T(A) = A^t$. Demuestre que T^2 es el operador identidad en $M_{n\times n}$.

14. Sea $T_1: \mathbf{R}^2 \to \mathbf{R}^3$ la transformación lineal

$$T_1(x, y) = (2x - y, x + 3y, 3x - 8y)$$

y $T_2 : \mathbf{R}^3 \to \mathbf{R}^3$ el operador lineal

$$T_2(x, y, z) = (x - y - z, 2x + 5y + z, 3x + 8y - 2z)$$

a) Al considerar las bases canónicas de $\mathbf{R}^2$ y $\mathbf{R}^3$, obtenga las matrices asociadas a T_1 y T_2.
b) Obtenga una fórmula explícita para $T_2 \circ T_1: \mathbf{R}^2 \to \mathbf{R}^3$.
c) Obtenga la matriz asociada a la transformación $T_2 \circ T_1$ del inciso anterior (respecto de las bases canónicas de $\mathbf{R}^2$ y $\mathbf{R}^3$) y compruebe que ésta es el producto de las matrices asociadas a T_2 y T_1 obtenidas en el inciso a).

15. Sea V un espacio vectorial de dimensión 3 y sea $\beta = \{v_1, v_2, v_3\}$ una base de V. Suponga que el operador $T_1: V \to V$ tiene asociada la siguiente matriz respecto de la base β:

$$A = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 0 & 3 \\ 1 & 1 & 1 \end{bmatrix}$$

y que el operador $T_2: V \to V$ tiene asociada la matriz (también respecto de la base β)

$$B = \begin{bmatrix} -1 & 1 & 0 \\ 3 & 2 & 1 \\ 1 & 4 & 2 \end{bmatrix}$$

Determine la matriz asociada a los operadores $T_1 \circ T_2$ y $T_2 \circ T_1$ respecto de la base β.

① 16. Suponga que el operador lineal $T_1: \mathbf{R}^2 \to \mathbf{R}^2$ tiene asociada la matriz

$$A = \begin{bmatrix} 1 & 3 \\ -2 & 1 \end{bmatrix}$$

respecto de la base $\beta = \{(3, 1), (2, 3)\}$ de $\mathbf{R}^2$, y el operador $T_2: \mathbf{R}^2 \to \mathbf{R}^2$ tiene asociada la matriz

$$B = \begin{bmatrix} -1 & 1 \\ 1 & 4 \end{bmatrix}$$

respecto de la base $\beta' = \{(1, 1), (-2, 4)\}$ de $\mathbf{R}^2$. Determine las matrices asociadas a los operadores $T_1 \circ T_2$ y $T_2 \circ T_1$ respecto de la base

a) β
b) β'
c) $\beta_1 = \{(2, 3), (3, 1)\}$
d) $\beta_2 = \{(1, 1), (2, 3)\}$
e) $\beta_3 = \{(0, 1), (1, 0)\}$
f) $\beta_4 = \{(1, 0), (0, 1)\}$

① 17. Considere las transformaciones lineales

$$T_1: V \to U \qquad T_2: U \to W \qquad T_2 \circ T_1: V \to W$$

Demuestre que

a) $Ker\, T_1 \subseteq Ker\, T_2 \circ T_1$
b) $Im\, T_2 \circ T_1 \subseteq Im\, T_2$

② 18. Sean T_1 y T_2 dos operadores lineales en el espacio vectorial V de dimensión finita, demuestre que

a) Nulidad $(T_2 \circ T_1) \geq$ máx(nulidad T_1, nulidad T_2)
b) Rango $(T_2 \circ T_1) \leq$ mín (rango T_1, rango T_2)

② 19. Sea $T: V \rightarrow V$ un operador lineal en el espacio vectorial V, compruebe que $Im\ T \cap Ker\ T = \{0\}$ si, y sólo si se tiene la siguiente implicación para todo $v \in V$

$$v \in Ker\, T^2 \Rightarrow v \in Ker\, T$$

② 20. Sea $T: V \rightarrow V$ un operador lineal en el espacio vectorial V de dimensión finita, demuestre que si rango T^2 = rango T, entonces
$Im\ T \cap Ker\ T = \{0\}$.

③ 21. Sea $T: V \rightarrow V$ un operador lineal en V tal que $T^2 = T$, pruebe que

a) $Ker\ T = Im\ (Id - T)$ (Id es el operador identidad en V)
b) $Ker\ (Id - T) = Im\ T$
c) $Ker\ T \cap Im\ T = \{0\}$
d) $V = Ker\ T \oplus Im\ T$

22. Verifique los resultados del ejercicio anterior con el operador lineal $T: \mathbf{R}^3 \rightarrow \mathbf{R}^3$ dado por

$$T(x, y, z) = (2x - y + z, -3x + 4y - 3z, -5x + 5y - 4z)$$

APÉNDICE (ÁLGEBRAS REALES)

23. Demuestre que el conjunto P de todos los polinomios es un álgebra conmutativa con unidad.

24. Sea A un álgebra real, el subconjunto no vacío $Ao \subseteq A$ es una *subálgebra* de A si Ao es por sí mismo un álgebra con las operaciones que estaban definidas en A (en particular, Ao es un subespacio del espacio vectorial A). Demuestre que el conjunto

$$A\circ = \{x \in A \mid xy = yx\ \forall\, y \in A\}$$

es una subálgebra de A. (A esta subálgebra se le llama "centro de A".)

25. Sea A un álgebra real y sea $x_o \in A$ un elemento fijo de A. Considere la función $T: A \rightarrow A$ dada por $T(y) = x_o y$. Demuestre que T es un operador lineal en el espacio vectorial A, esto es, demuestre que

$$T(y_1 + y_2) = T(y_1) + T(y_2)$$
$$T(cy) = cT(y) \qquad \forall\, y, y_1, y_2 \in A \quad c \in \mathbf{R}$$

② 26. Sea A un álgebra real con unidad. Considere el álgebra $L(A, A)$ de todos los operadores lineales en A. Sea $F: A \rightarrow L(A, A)$ la función dada por $F(x) = T_x$ en donde $T_x: A \rightarrow A$ es el operador lineal $T_x(y) = xy$ (véase ejercicio anterior). Pruebe que para todo $x, y \in A$ y todo $c \in \mathbf{R}$ se tiene

a) $F(x + y) = F(x) + F(y)$
b) $F(cx) = cF(x)$
c) $F(xy) = F(x)\, F(y)$
d) F es inyectiva

④ 27. (Un ejercicio sobre álgebras de Lie.) Sea V un espacio vectorial. Si en V está definida una "operación corchete" [,]: $V \times V \rightarrow V$ con las propiedades siguientes:

1) $[x, y] = -[y, x]$
2) $[ax + by, z] = a[x, z] + b[y, z], a, b, \quad \mathbf{R}$
3) $[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$

se dice que V es un *álgebra de Lie.*

a) Considere el espacio vectorial $\mathbf{R}^3$. Para $X, Y \in \mathbf{R}^3$ defina el corchete de X y Y como $[X, Y] = X \times Y$ (el producto vectorial o producto cruz de X y Y), es decir, si $X = (x_1\ x_2, x_3)$ y $Y = (y_1, y_2, y_3)$

$$X \times Y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)$$

Demuestre que $\mathbf{R}^3$ con esta operación corchete es un álgebra de Lie.

b) Considere el espacio vectorial $M_{n \times n}$. Sean A y B dos matrices de este espacio, defina el corchete de A y B como

$$[A, B] = AB - BA$$

Compruebe que $M_{n \times n}$ con esta operación corchete es un álgebra de Lie.

c) Considere el espacio vectorial $L(V, V)$ de los operadores lineales en el espacio vectorial V. Sean T_1 y T_2 dos operadores de este espacio, defina el corchete de T_1 y T_2 como

$$[T_1, T_2] = T_1 \circ T_2 - T_2 \circ T_1$$

Demuestre que $L(V, V)$ con esta operación corchete es un álgebra de Lie.

d) Sean V_1 y V_2 dos álgebras de Lie, se dice que el álgebra V_1 es *isomorfa* al álgebra V_2 si existe un isomorfismo de espacios vectoriales $T: V_1 \rightarrow V_2$ tal que

$$[Tx, Ty]_2 = T([x, y]_1) \qquad \forall\, x, y \in V_1$$

en donde $[\ ,\ ]_1$ y $[\ ,\ ]_2$ son los corchetes de V_1 y V_2, respectivamente. Compruebe que el álgebra de Lie $M_{n \times n}$ es isomorfa al álgebra de Lie $L(V, V)$ en donde V es un espacio vectorial de dimensión n.

6. INVERSA DE UNA TRANSFORMACIÓN LINEAL

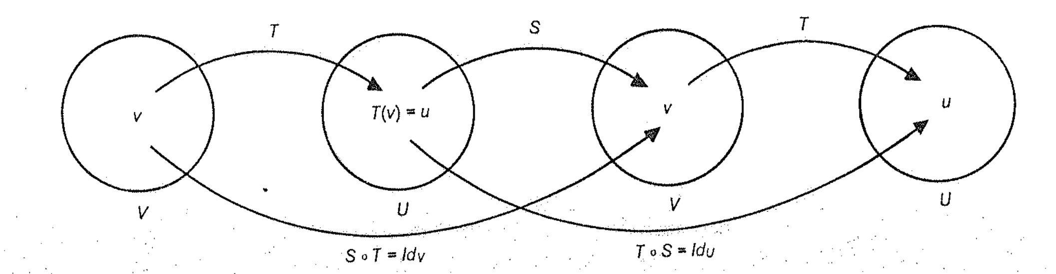
En esta sección se aborda el siguiente problema: dada una transformación lineal $T: V \rightarrow U$ del espacio V al espacio U, ¿bajo qué condiciones se puede hablar de "una inversa" de T? (En el sentido de funciones, por ejemplo.) El objetivo es dar condiciones equivalentes al hecho de que la trans formación T sea inversible (en el sentido preciso de la definición que se dará a continuación), estudiar las

propiedades de la inversa de T, cuando ésta existe y, por último, relacionar este análisis con las matrices que representan a las transformaciones correspondientes (en el caso de ser V y U espacios de dimensión finita).

Algunas de las ideas que aparecerán en esta sección ya fueron presentadas en el capítulo anterior, cuando se habló de isomorfismos (subsección 6.2). Sin embargo, para presentar una continuidad en el análisis que ahora se emprenderá, tales ideas se expondrán de nuevo, en su momento.

DEFINICIÓN 6.1 Sean V y U espacios vectoriales y $T: V \to U$ una transformación lineal. Se dice que la transformación T es *invertible* si existe una transformación lineal $S: U \to V$ tal que $S \circ T = Id_V$ y $T \circ S = Id_U$, en donde Id_V e Id_U son los operadores identidad en V y U respectivamente [esto es, $Id_V(v) = v \ \forall\, v \in V$ e $Id_U(u) = u \ \forall\, u \in U$].

Esquemáticamente



Obsérvese que en la definición anterior se pide que para que $T: V \to U$ sea inversible, debe existir $S: U \to V$ tal que $S \circ T = Id_V$ y $T \circ S = Id_U$. Aun cuando se tenga un operador lineal $T: V \to V$, el hecho de que exista $S: V \to V$ tal que $S \circ T = Id_V$, esto *no garantiza* que $T \circ S = Id_U$ pues, como se sabe, la composición de transformaciones lineales es una operación no conmutativa.

Por ejemplo, si P es el espacio vectorial de *todos* los polinomios, considérese el operador lineal $T: P \to P$, dado por $T(p) = p'$(la derivada de p). Sea $S: P \to P$ el operador

$$S(a_0 + a_1x + \ldots + a_kx^k) = a_0x + a_1\frac{x^2}{2} + \ldots + a_k\frac{x^{k+1}}{k+1}$$

Es fácil verificar que S es lineal.

Obsérvese que

$$(T \circ S)(a_0 + a_1x + \ldots + a_kx^k) = T\left(a_0x + a_1\frac{x^2}{2} + \ldots + a_k\frac{x^{k+1}}{k+1}\right)$$
$$= a_0 + a_1x + \ldots + a_kx^k$$

de modo que $T \circ S = Id_p$.

Sin embargo, si $a_o \neq 0$

$$(S \circ T)(a_0 + a_1x + \ldots + a_kx^k) = S(a_1 + 2a_2x + \ldots + ka_kx^{k-1})$$
$$= a_1x + a_2x^2 + \ldots + a_kx^k$$
$$\neq a_0 + a_1x + a_2x_2 + \ldots + a_kx^k$$

de modo que $S \circ T \neq Id_p$, y entonces T no es inversible. Se puede demostrar (véanse ejercicios al final de esta sección) que situaciones como la que muestra el ejemplo anterior no pueden ocurrir si V es un espacio de dimensión finita.

Antes de comenzar a desarrollar la teoría sobre inversibilidad de transformaciones lineales, véase el siguiente ejemplo:

EJEMPLO 1 Sea $T: \mathbf{R}^2 \to P_1$ la transformación $T(a, b) = a + b + (2a + 3b)x$. Se verifica fácilmente que T es lineal. Uno se pregunta si T es inversible.

Proceda directamente con la definición 6.1 para contestar esta pregunta.

Si T fuera inversible, debería existir una transformación lineal $S: P_1 \to \mathbf{R}^2$ tal que

$$S \circ T = Id_{\mathbf{R}^2}$$

y

$$T \circ S = Id_{P_1}$$

Al usar la primera de estas relaciones vea cómo tendría que ser la transformación S.

Se tiene

$$(a, b) = Id_{\mathbf{R}^2}(a, b) = (S \circ T)(a, b) = S(T(a, b)) = S(a + b + (2a + 3b)x)$$

Al escribir $a + b = \gamma$ y $2a + 3b = \delta$, se puede resolver para a y b en términos de γ y δ y obtener

$$a = 3\gamma - \delta \qquad b + \delta - 2\gamma$$

Entonces la transformación $S: P_1 \to \mathbf{R}^2$ tendría que ser

$$S(\gamma + \delta x) = (3\gamma - \delta, \delta - 2\gamma)$$

Se puede verificar que en efecto S es una transformación lineal y que $T \circ S = Id_{P_1}$. Se concluye entonces que T es inversible.

Una primera observación importante que se debe hacer de la definición 6.1 es que si T es una transformación lineal inversible, la transformación lineal S es única.

En efecto, suponga que existen S_1 y S_2 satisfaciendo las condiciones de la definición 6.1. Entonces

$$S_1 = Id_V \circ S_1 = (S_2 \circ T) \circ S_1 = S_2 \circ (T \circ S_1) = S_2 \circ Id_U = S_2$$

Por lo tanto, si la transformación lineal $T: V \to U$ es inversible, se puede definir a la transformación S como *la* inversa de T, la cual se denotará por T^{-1}.

Otro par de consecuencias inmediatas de la definición 6.1, se encuentra en el siguiente teorema:

TEOREMA 6.1

1) Si la transformación lineal $T: V \to U$ es inversible entonces $T^{-1}: U \to V$ también es inversible y $(T^{-1})^{-1} = T$.
2) Si $T_1: V \to U$ y $T_2: U \to W$ son transformaciones lineales inversibles, entonces $T_2 \circ T_1: V \to W$ también es inversible y $(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1}$.

DEMOSTRACIÓN 1) Es una consecuencia directa de la definición 6.1 y de la unicidad de la inversa de una transformación. Para verificar 2) basta observar que $T_1^{-1} \circ T_2^{-1}$ es lineal (composición de transformaciones lineales) y además

$$(T_1^{-1} \circ T_2^{-1}) \circ (T_2 \circ T_1) = T_1^{-1} \circ (T_2^{-1} \circ T_2) \circ T_1 = T_1^{-1} \circ (Id_V) \circ T_1 = T_1^{-1} \circ T_1 = Id_V$$

$$(T_2 \circ T_1) \circ (T_1^{-1} \circ T_2^{-1}) = T_2 \circ (T_1 \circ T_1^{-1}) \circ T_2^{-1} = T_2 \circ (Id_U) \circ T_2^{-1} = T_2 \circ T_2^{-1} = Id_W$$

y por tanto, $(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1}$.

Q.E.D.

El próximo objetivo es caracterizar la inversibilidad de una transformación lineal en términos de su inyectividad y de su sobreyectividad. Se verá en el teorema 6.2 que la transformación T es inversible si, y sólo si es inyectiva y sobreyectiva. Pero antes véase el siguiente lema que caracteriza la propiedad de inyectividad de una transformación lineal en términos de su núcleo.

LEMA

Sea $T: V \to U$ una **transformación lineal**, entonces T es inyectiva si, y sólo sí $Ker\ T = \{0\}$.

DEMOSTRACIÓN Supóngase que T es inyectiva y sea v un vector del núcleo de T. Entonces $T(v) = 0$, pero también $T(0) = 0$ (por ser T lineal). Entonces $v = 0$, esto es, $Ker\ T = \{0\}$. Recíprocamente, supóngase que $Ker\ T = \{0\}$, y sean $v, v' \in V$ tales que $T(v) = T(v')$. Entonces $T(v) - T(v') = T(v - v') = 0$, lo que dice que $v - v' \in Ker\ T$. Por tanto, $v - v' = 0$, esto es, $v = v'$. Esto muestra que T es inyectiva.

Q.E.D.

TEOREMA 6.2

La transformación lineal $T: V \to U$ es inversible si, y sólo si T es inyectiva y sobreyectiva.

DEMOSTRACIÓN Supóngase primeramente que T es inversible. Sea v un vector del núcleo T. Entonces

$$T^{-1}(T(v)) = T^{-1}(0) = 0 \qquad \text{(pues } T^{-1} \text{ es lineal)}$$

Pero $T^{-1} \circ T = Id_V$. Es decir,

$$v = (T^{-1} \circ T)(v) = T^{-1}(T(v)) = 0$$

Por lo tanto, $Ker\ T = \{0\}$ y entonces, según el lema anterior, T es inyectiva.

Sea ahora $u \in U$. Se debe mostrar que existe $v \in V$ tal que $T(v) = u$ (esto es, que $Im\ T = U$). Como la transformación T^{-1} está definida en U, escríbase $v = T^{-1}(u)$.

Entonces

$$T(v) = T(T^{-1}(u)) = (T \circ T^{-1})(u) = Id_U(u) = u$$

lo que muestra entonces que T es sobreyectiva.

Recíprocamente, supóngase que T es inyectiva y sobreyectiva. Se debe mostrar que existe una transformación lineal $T^{-1}: U \to V$ tal que $T^{-1} \circ T = Id_V$ y $T \circ T^{-1} = Id_U$.

Para cada vector $u \in U$, existe un vector $v \in V$ tal que $T(v) = u$, pues T es sobreyectiva. Además este vector v es único, pues T es inyectiva. Defínase $T^{-1}: U \to V$ como $T^{-1}(u) = v$. Es obvio que se satisfacen $T^{-1} \circ T = Id_V$ y $T \circ T^{-1} = Id_U$. Sólo resta probar que T^{-1} es una transformación lineal.

Sean u y u' dos vectores cualesquiera de U y considérense los correspondientes vectores v y v' en V tales que $T(v) = u$, $T(v') = u'$. Entonces $u + u' = T(v) + T(v') = T(v + v')$ pues T es lineal, de donde

$$T^{-1}(u + u') = v + v' = T^{-1}(u) + T^{-1}(u')$$

Si $c \in \mathbf{R}$, se tiene que $T(cv) = cT(v) = cu$, por lo que $T^{-1}(cu) = cv = cT^{-1}(u)$.

Esto muestra entonces que T^{-1} es lineal, y por lo tanto T es inversible.

Q.E.D.

En el capítulo anterior, se definió el concepto de *isomorfismo* como una transformación lineal biyectiva (inyectiva y sobreyectiva). Con el lenguaje de esta sección se tiene entonces la siguiente definición (consecuencia del teorema anterior).

DEFINICIÓN 6.2 Un *isomorfismo* es una transformación lineal inversible.

También en el capítulo anterior se probó que dos espacios vectoriales V y U con la misma dimensión (finita) son isomorfos (es decir, existe un isomorfismo $T: V \to U$). Ahora se está en posibilidades de demostrar una especie de recíproco de esta afirmación.

TEOREMA 6.3

Supóngase que V es un espacio vectorial isomorfo al espacio vectorial U. Entonces V es de dimensión finita si, y sólo si U es de dimensión finita (y $dim\ U = dim\ V$).

DEMOSTRACIÓN Sea $T: V \to U$ un isomorfismo de V a U. Supóngase que V es de dimensión finita. Se puede entonces aplicar el teorema de la dimensión (teorema 2.2) que dice: rango de T + nulidad de $T = dim\ V$. Siendo T inyectiva, $Ker\ T = \{0\}$ y por tanto nulidad de $T = 0$. Entonces rango de $T = dim\ V$. Pero por otra parte, T es sobreyectiva, de modo que $Im\ T = U$. Entonces rango de $T = dim\ Im\ T = dim\ U = dim\ V$ (finita). La implicación recíproca se prueba similarmente considerando el isomorfismo inverso $T^{-1}: U \to V$.

Q.E.D.

COROLARIO 1

Dos espacios vectoriales de dimensión finita son isomorfos si, y sólo si tienen la misma dimensión.

DEMOSTRACIÓN Por lo probado en el capítulo anterior y en el teorema 6.3.

Q.E.D.

COROLARIO 2

Sea V un espacio vectorial de dimensión finita y sea $T: V \to U$ una transformación lineal. Una condición necesaria (no suficiente) para que la transformación T sea inversible es que U sea de dimensión finita y que $dim\ U = dim\ V$.

DEMOSTRACIÓN Se sigue inmediatamente del teorema 6.3.

Q.E.D.

NOTA: La condición $dim\ U = dim\ V$ del corolario anterior no garantiza en absoluto que la transformación $T: V \to U$ sea inversible. Lo que se sabe es que si $dim\ U = dim\ V$, *existe una* transformación lineal inversible $S: V \to U$ (que no tiene por qué coincidir con T).

Retómese el análisis sobre la inversibilidad de una transformación lineal $T: V \to U$. Según el teorema 6.2, esta transformación puede no ser inversible por dos causas (independientes): T no es inyectiva y/o T no es sobreyectiva. Si T no es inyectiva, no hay manera alguna de definir una transformación inversa $T^{-1}: U \to V$, pues en este caso existirían dos vectores $v_1, v_2 \in V$ tales que $T(v_1) = T(v_2) = u_o \in U$, de modo que $T^{-1}(u_o)$ no estaría bien definido.

Sin embargo, si T no es sobreyectiva, pero sí es inyectiva, se podría modificar su codominio de modo que T (con su codominio redefinido) sea inversible. Basta considerar el codominio de T como siendo su imagen (que se sabe que es un espacio vectorial, subespacio de U) y entonces la transformación $T': V \to Im\ T$ $T'(v) = T(v)\ \forall\ v \in V$ es ya inversible (pues ahora claramente es sobreyectiva). Las transformaciones lineales que son inyectivas (independientemente si son sobreyectivas o no lo son) reciben un nombre especial.

DEFINICIÓN 6.3 Se dice que la transformación lineal $T: V \to U$ es *no singular* si T es inyectiva (o equivalente si $Ker\ T = \{0\}$).

EJEMPLO 2 Entonces, toda transformación lineal inversible es no singular. La afirmación recíproca es (obviamente) falsa, como lo muestra el siguiente ejemplo:

Sea $T: \mathbf{R}^2 \to \mathbf{R}^3$ la transformación lineal $T(x, y) = (x, y, x)$. Obsérvese que si $(x, y, x) = 0 \Rightarrow x = y = 0$, de modo que $Ker\ T = \{0\}$. Entonces T es no singular.

Sin embargo, T no puede ser inversible pues $dim\ \mathbf{R}^2 = 2 \neq 3 = dim\ \mathbf{R}^3$ (corolario 2 del teorema 6.3).

En el caso particular en el que V es un espacio vectorial de dimensión finita, existen una serie de propiedades que caracterizan la inversibilidad de la transformación $T: V \to U$. Estas propiedades se enuncian en el siguiente teorema, cuya demostración ya se había prometido anteriormente (subsección 6.2, capítulo 3 y subsección 3.1 del presente capítulo). Obsérvese que en este caso se tiene necesariamente $dim\ U = dim\ V$ (corolario 2 del teorema 6.3).

TEOREMA 6.4

Sean V y U espacios vectoriales de dimensión finita tales que $dim\ V = dim\ U$. Las siguientes afirmaciones acerca de la transformación lineal $T: V \to U$ son equivalentes

1) T es inversible (esto es, T es un isomorfismo)
2) T es sobreyectiva
3) T es inyectiva
4) T es no singular
5) Si $\{v_1, v_2, \ldots, v_k\}$ es un conjunto de vectores linealmente independientes en V, entonces $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ es un conjunto de vectores linealmente independientes en U.
6) Si $\{v_1, v_2, \ldots, v_n\}$ es una base de V, entonces $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ es una base de U.

DEMOSTRACIÓN Se demostrará (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (6) $\Rightarrow$ (1).

(1) $\Rightarrow$ (2). Obvio (teorema 6.2)

(2) $\Rightarrow$ (3). Por hipótesis T es sobreyectiva. Entonces $Im\ T = U$ y rango de T =

dim Im T = *dim* U = *dim* V. Por el teorema de la dimensión se tiene nulidad de $T = 0$. Entonces *Ker* $T = \{0\}$. Entonces T es inyectiva (lema previo al teorema 6.2).

(3) $\Rightarrow$ (4). Obvio (por el lema previo al teorema 6.2).

(4) $\Rightarrow$ (5). Sea entonces $\{v_1, v_2, \ldots, v_k\}$ un conjunto de vectores *linealmente independientes* en V. Considérese la combinación lineal

$$c_1T(v_1) + c_2T(v_2) + \ldots + c_kT(v_k) = 0$$

o sea

$$T(c_1v_1 + c_2v_2 + \ldots + c_kv_k) = 0$$

Como T es no singular, la relación anterior implica que

$$c_1v_1 + c_2v_2 + \ldots + c_kv_k = 0$$

la cual a su vez implica que $c_1 = c_2 = \ldots = c_k = 0$ (por la independencia lineal de los vectores $v_1, v_2, \ldots, v_k$) lo que muestra que los vectores $T(v_1), T(v_2), \ldots, T(v_k)$ son *linealmente independientes*.

(5) $\Rightarrow$ (6). Sea $\{v_1, v_2, \ldots, v_n\}$ una base de V. (Entonces n = *dim* V = *dim* U). El conjunto de n vectores $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ es un conjunto linealmente independiente en U (por hipótesis). Como *dim* $U = n$, este conjunto de hecho forma una base de U (corolario del teorema 5.3, capítulo 3).

(6) $\Rightarrow$ (1). Sea $\{v_1, v_2, \ldots, v_n\}$ una base de V. Por hipótesis se sabe que $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ es una base de U. Sea $u \in U$. Entonces existen escalares $c_1, c_2, \ldots, c_n$ tales que $u = c_1T(v_1) + c_2T(v_2) + \ldots + c_nT(v_n) = T(c_1v_1 + c_2v_2 + \ldots + c_nv_n)$. Escríbase $v = c_1v_1 + c_2v_2 + \ldots + c_nv_n$. Se ha probado así que dado $u \in U$, existe $v \in V$ tal que $T(v) = u$, es decir, que T es sobreyectiva. Por otra parte, supóngase que $v = d_1v_1 + d_2v_2 + \ldots + d_nv_n$ es un vector del núcleo de T. Entonces $T(v) = T(d_1v_1 + d_2v_2 + \ldots + d_nv_n) = d_1T(v_1) + d_2T(v_2) + \ldots + d_nT(v_n) = 0$, relación que implica que $d_1 = d_2 = \ldots = d_n = 0$ (pues los vectores $T(v_1), T(v_2), \ldots, T(v_n)$ son linealmente independientes), o sea que $v = 0$. Se ha probado entonces que *Ker* $T = \{0\}$, y por tanto T es inyectiva. Esto muestra que T es un isomorfismo.

Q.E.D.

El teorema anterior brinda varias alternativas equivalentes para verificar si una transformación lineal entre espacios vectoriales de la misma dimensión es o no inversible. Sin embargo, algunas de ellas son claramente poco prácticas.

A continuación se comenzará el estudio de la inversibilidad de una transformación lineal $T: V \to U$ (en donde *dim* V = *dim* U) vía su representación matricial respecto de bases concretas de V y U. Este estudio reportara "condiciones numéricas" (fáciles de evaluar) equivalentes a la inversibilidad de la transformación T.

El resultado principal se encuentra en el siguiente teorema:

TEOREMA 6.5

Sean V y U espacios vectoriales de la misma dimensión (finita), la transformación lineal $T: V \to U$ es inversible si, y sólo si la matriz de la transformación T respecto de cualesquiera bases de V y U, es una matriz inversible.

DEMOSTRACIÓN Supóngase que $T: V \to U$ es inversible. Existe entonces una transformación lineal $S: U \to V$ tal que $S \circ T = Id_V$ y $T \circ S = Id_U$. Sean β_1 y β_2 bases de V y U respectivamente, obsérvese que $[Id_V]_{\beta_1} = I$ e $[Id_U]_{\beta_2} = I$. Entonces, según la fórmula 5.2 se tiene que

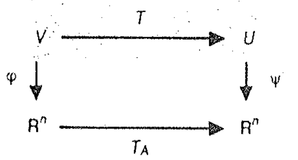
$$I = [Id]_{\beta_1} = [S \circ T]_{\beta_1} = [S]_{\beta_2\beta_1}\,[T]_{\beta_1\beta_2}$$

y

$$I = [Id]_{\beta_2} = [T \circ S]_{\beta_2} = [T]_{\beta_1\beta_2}\,[S]_{\beta_2\beta_1}$$

lo cual dice entonces que $[T]_{\beta_1\beta_2}$ es una matriz inversible.

Supóngase ahora que la matriz $A = [T]_{\beta_1\beta_2}$ es inversible. Considérese la transformación $T_A: \mathbf{R}^n \to \mathbf{R}^n$, n = *dim* V = *dim* U, dada por $T_A(X) = AX$ y sean $\varphi: V \to \mathbf{R}^n$ y $\psi: U \to \mathbf{R}^n$ los isomorfismos naturales $\varphi(v) = (v)_{\beta_1}$ y $\psi(u) = (u)_{\beta_2}$. Se tiene entonces una situación como la siguiente:



Si el vector $X \in \mathbf{R}^n$ pertenece al núcleo de T_A, entonces $T_A(X) = AX = 0$. Siendo A inversible, el sistema homogéneo de ecuaciones lineales $AX = 0$ tiene sólo la solución trivial $X = 0$. Entonces *Ker* $T_A = \{0\}$. Por otra parte, se sabe que $\varphi($*Ker* $T)$ = *Ker* T_A. Entonces $\varphi($*Ker* $T) = \{0\}$ y por tanto *Ker* $T = \{0\}$. Esto dice, según el teorema 6.4 que T es inversible.

Q.E.D.

COROLARIO

(DE LA DEMOSTRACIÓN ANTERIOR.) Si la transformación $T: V \to U$ es inversible, entonces

$$[T^{-1}]_{\beta_2\beta_1} = [T]^{-1}_{\beta_1\beta_2}$$

en donde β_1 y β_2 son bases de V y U, respectivamente.

DEMOSTRACIÓN Es una consecuencia de las fórmulas

$$[S]_{\beta_2\beta_1}\,[T]_{\beta_1\beta_2} = [T]_{\beta_1\beta_2}\,[S]_{\beta_2\beta_1} = I$$

obtenidas en la demostración del teorema anterior, en donde $S = T^{-1}$.

Q.E.D.

EJEMPLO 3 Véase de nuevo el primer ejemplo que se consideró en esta sección a la luz del teorema anterior. Se tenía la transformación lineal

$$T: \mathbf{R}^2 \to P_1 \qquad T(a, b) = a + b + (2a + 3b)x$$

Sean β_1 y β_2 las bases canónicas de $\mathbf{R}^2$ y P_1, respectivamente. Halle la matriz $[T]_{\beta_1\beta_2}$.

$$T(1, 0) = 1 + 2x \qquad [T(1, 0)]_{\beta_2} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$T(0, 1) = 1 + 3x \qquad [T(0, 1)]_{\beta_2} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Entonces

$$[T]_{\beta_1\beta_2} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$$

Como $det\ [T]_{\beta_1\beta_2} = 1 \neq 0$, se concluye que T es inversible (pues la matriz $[T]_{\beta_1\beta_2}$ lo es). Más aún, si se halla $[T]^{-1}_{\beta_1\beta_2}$, ésta será la matriz de la transformación inversa $T^{-1}: P_1 \to \mathbf{R}^2$ respecto de las bases β_2 y β_1.

Se tiene

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{array}\right] \sim \cdots \sim \left[\begin{array}{cc|cc} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 1 \end{array}\right]$$

de modo que

$$[T]^{-1}_{\beta_1\beta_2} = [T^{-1}]_{\beta_2\beta_1} = \begin{bmatrix} 3 & -1 \\ -2 & 1 \end{bmatrix}$$

y entonces si $p = \gamma + \delta x$, se tiene

$$[T^{-1}(p)]_{\beta_1} = [T^{-1}]_{\beta_2\beta_1}[p]_{\beta_2} = \begin{bmatrix} 3 & -1 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} \gamma \\ \delta \end{bmatrix} = \begin{bmatrix} 3\gamma - \delta \\ -2\gamma + \delta \end{bmatrix}$$

o sea

$$T^{-1}(\gamma + \delta x) = (3\gamma - \delta, -2\gamma + \delta)$$

resultado al que se había llegado ya anteriormente.

Considere ahora el caso particular de un *operador* lineal $T: V \to V$ en donde V es de dimensión finita. Sea β una base cualquiera de V.

DEFINICIÓN 6.4 Sea $T: V \to V$ un operador lineal en el espacio vectorial V de dimensión finita. Se llama *determinante de T*, denotado por *det T*, al determinante de la matriz que representa a T en cualquier base de V.

Verifíquese primeramente que esta definición es una buena definición, en el sentido de que el valor *det T no depende* de la base que se escoja del espacio V.

Sean β y β' dos bases distintas de V y considerándose las matrices $[T]_\beta$ y $[T]_{\beta'}$. Según el corolario del teorema 4.1, estas matrices son semejantes. Existe entonces una matriz inversible P tal que

$$[T]_\beta = P^{-1}[T]_{\beta'}P$$

(de hecho se sabe que P es la matriz de cambio de base de β a β'). Entonces:

$$det[T]_\beta = det(P^{-1}[T]_{\beta'}P) = (det\ P^{-1})(det[T]_{\beta'})(det\ P)$$
$$= (det P)^{-1}(det[T]_{\beta'})(det\ P) = det[T]_{\beta'}$$

lo que muestra entonces que el valor de *det T* no depende de la base que se escoja de V.

Se ha asociado entonces a cada operador lineal $T: V \to V$ un número bien determinado: *det T*. Este número dice si el operador es inversible, como lo muestra el siguiente teorema:

TEOREMA 6.6 El operador lineal $T: V \to V$, donde V es un espacio vectorial de dimensión finita, es inversible si, y sólo si $det\ T \neq 0$.

DEMOSTRACIÓN Consecuencia directa del teorema 6.5 y de la definición de *det T*.

Q.E.D.

Por ejemplo, considérese el operador $T: M_{2\times 2} \to M_{2\times 2}$ dado por

$$T\left(\begin{bmatrix} a & b \\ c & c \end{bmatrix}\right) = \begin{bmatrix} 5a + b & 2a - b + c \\ 3c + 2d & a + b + 2d \end{bmatrix}$$

Tómese la base canónica de $M_{2\times 2}$ β.

Entonces

$$[T]_\beta = \begin{bmatrix} 5 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ 0 & 0 & 3 & 2 \\ 1 & 1 & 0 & 2 \end{bmatrix}$$

Como $det\ [T]_\beta = -34 \neq 0$, se concluye que T es inversible.

Para terminar esta sección, se va a considerar un problema "elemental" que se atacará con las ideas que se han desarrollado en este capítulo. El objetivo es reobtener, a la luz de la teoría de transformaciones lineales estudiada hasta este momento, el teorema 6.1 del capítulo 3, en el que se resumía la discusión hecha en la subsección 6.1 de tal capítulo sobre cambio de bases.

Considérese el operador identidad en el espacio vectorial V de dimensión finita $Id: V \to V$, $Id(v) = v \ \forall\ v \in V$. Se van a considerar dos bases distintas de V, a saber: la base β_1 en su dominio y la base β_2 en su codominio. Escríbase $\beta_1 = \{v_1, v_2, \ldots, v_n\}$ y $\beta_2 = \{v_1', v_2', \ldots, v_n'\}$. Obsérvese que la matriz de este operador respecto de las bases β_1 y β_2 es (esquemáticamente)

$$[Id]_{\beta_1\beta_2} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ [v_1]_{\beta_2} & [v_2]_{\beta_2} & \cdots & [v_n]_{\beta_2} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

que es precisamente la matriz que se definió en la subsección 6.1 del capítulo 3 como "matriz de cambio de base de β_1 a β_2", y que se denotó por Q.

Se sabe que (sección 3), para $v \in V$

$$[Id(v)]_{\beta_2} = [v]_{\beta_2} = [Id]_{\beta_1\beta_2}[v]_{\beta_1}$$

o sea

$$[v]_{\beta_2} = Q[v]_{\beta_1}$$

(fórmula que ya conocíamos del capítulo 3).

Además, obviamente el operador Id es inversible, de modo que $Id^{-1} = Id$ tiene por matriz respecto de las bases β_2 y β_1.

$$[Id^{-1}]_{\beta_2\beta_1} = [Id]^{-1}_{\beta_1\beta_2} = Q^{-1}$$

Es claro que $[Id^{-1}]_{\beta_2\beta_1} = [Id]_{\beta_2\beta_1}$, es la matriz de cambio de base de β_2 a β_1, la cual se denotaba como P.

Éstos son precisamente los resultados que se obtuvieron en la subsección 6.1 del capítulo 3. Claro que en aquel momento tomó mucho más trabajo llegar a ellos.

No se puede negar que, en contraste, la manera como se han establecido aquí tales resultados es una manera sumamente simple y elegante.

APÉNDICE. UNA APLICACIÓN AL CÁLCULO DE ANTIDERIVADAS

En todo este apéndice, V denotará un subespacio de dimensión finita del espacio vectorial de las funciones reales con derivadas de todos los órdenes.

Considérese el operador $D: V \to V$ dado por $D(f) = f'$ (la derivada de f).

Como ya se había observado en la sección 1 de este capítulo, D es un operador lineal pues

$$D(f+g) = (f+g)' = f' + g' = D(f) + D(g) \qquad f, g \in V$$

$$D(cf) = (cf)' = cf' = cD(f) \qquad c \in \mathbf{R}$$

Se van a considerar algunos casos particulares en los cuales, el operador D es inversible.

Sea $\beta = \{f_1, f_2, \ldots, f_n\}$ una base de V y sea A la matriz que representa a D respecto de la base β. Si $g \in V$, entonces, como se sabe

$$[D(g)]_\beta = A[g]_\beta \qquad \text{(A1)}$$

Siendo D inversible, se sabe también que el operador $D^{-1}: V \to V$ es representado respecto de la base β por la matriz A^{-1}. Es decir,

$$[D^{-1}(g)]_\beta = A^{-1}[g]_\beta \qquad \text{(A2)}$$

Obsérvese que $D^{-1}(g)$ es la función en V cuya imagen bajo D es g. En otras palabras $D^{-1}(g)$ es la función cuya derivada es g. Entonces

$$D^{-1}(g) = \int g(x)\, dx$$

(la antiderivada de g).

Según la fórmula (A2), se tiene entonces que para hallar la antiderivada de una función $g \in V$ (su matriz de coordenadas respecto de la base β), sólo se tiene que determinar la inversa de la matriz A. Éste es entonces un "método", que emplea herramientas del álgebra lineal, para hallar antiderivadas de funciones.

Véanse funcionar estas ideas en algunos casos concretos.

Sea V el espacio generado por las funciones sen ax, cos ax, x sen ax, x cos ax. Se puede demostrar que estas funciones son linealmente independientes. Entonces $\beta = \{\text{sen } ax, \cos ax, x \text{ sen } ax, x \cos ax\}$ es una base de V. Hállese la matriz de D respecto de la base β. Se tiene

$$D(\text{sen } ax) = a \cos ax$$

$$D(\cos ax) = -a \text{ sen } ax$$

$$D(x \text{ sen } ax) = \text{sen } ax + ax \cos ax$$

$$D(x \cos ax) = \cos ax - ax \text{ sen } ax$$

y entonces

$$(D(\text{sen } ax))_\beta = (0, a, 0, 0)$$

$$(D(\cos ax))_\beta = (-a, 0, 0, 0)$$

$$(D(x \text{ sen } ax))_\beta = (1, 0, 0, a)$$

$$(D(x \cos ax))_\beta = (0, 1, -a, 0)$$

La matriz A es pues

$$A = \begin{bmatrix} 0 & -a & 1 & 0 \\ a & 0 & 0 & 1 \\ 0 & 0 & 0 & -a \\ 0 & 0 & a & 0 \end{bmatrix}$$

cuya inversa es la matriz

$$A^{-1} = \begin{bmatrix} 0 & 1/a & 1/a^2 & 0 \\ -1/a & 0 & 0 & 1/a^2 \\ 0 & 0 & 0 & 1/a \\ 0 & 0 & -1/a & 0 \end{bmatrix}$$

Según la fórmula (A2) se tiene entonces que

$$[D^{-1}(\text{sen } ax)]_\beta = A^{-1}[\text{sen } ax]_\beta = \begin{bmatrix} 0 & 1/a & 1/a^2 & 0 \\ -1/a & 0 & 0 & 1/a^2 \\ 0 & 0 & 0 & 1/a \\ 0 & 0 & -1/a & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1/a \\ 0 \\ 0 \end{bmatrix}$$

o sea que

$$D^{-1}(\text{sen } ax) = \int \text{sen } ax\, dx = -\frac{1}{a} \cos ax$$

De la misma manera se obtienen las fórmulas

$$\int \cos ax\, dx = \frac{1}{a} \text{sen } ax$$

$$\int x \text{ sen } ax\, dx = \frac{1}{a^2} \text{sen } ax - \frac{1}{a} x \cos ax$$

$$\int x \cos ax\, dx = \frac{1}{a^2} \cos ax - \frac{1}{a} x \text{ sen } ax$$

Supóngase que se quiere calcular la integral

$$J = \int (3 \text{ sen } ax - 2 \cos ax + 5x \text{ sen } ax - 4x \cos ax)\, dx$$

Entonces $J = D^{-1}(g)$, en donde $(g)_\beta = (3, -2, 5, -4)$. Por tanto, según la fórmula (A2) se tiene que

$$[J]_\beta = A^{-1}[g]_\beta = \begin{bmatrix} 0 & 1/a & 1/a^2 & 0 \\ -1/a & 0 & 0 & 1/a^2 \\ 0 & 0 & 0 & 1/a \\ 0 & 0 & -1/a & 0 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 5 \\ -4 \end{bmatrix} = \begin{bmatrix} -2/a + 5/a^2 \\ -3/a - 4/a^2 \\ -4/a \\ -5/a \end{bmatrix}$$

y así entonces

$$J = (2/a + 5/a^2) \text{ sen } ax + (-3/a - 4/a^2) \cos ax - 4/a\, x \text{ sen } ax - 5/a\, x \cos ax$$

Como siguiente ejemplo, considérese el espacio V generado por las funciones e^{ax} sen bx y $e^{ax}\cos bx$. Nuevamente, es fácil probar que estas funciones son linealmente independientes, de modo que ellas constituyen una base β de V.

Determínese la matriz de D respecto de la base β. Se tiene que

$$D(e^{ax} \text{ sen } bx) = ae^{ax} \text{ sen } bx + be^{ax} \cos bx$$

$$D(e^{ax} \cos bx) = -be^{ax} \text{ sen } bx + ae^{ax} \cos bx$$

y entonces

$$(D(e^{ax} \text{ sen } bx))_\beta = (a, b)$$

$$(D(e^{ax} \cos bx))_\beta = (-b, a)$$

así que la matriz A de D respecto de la base β es

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

y su inversa

$$A^{-1} = \begin{bmatrix} \dfrac{a}{a^2+b^2} & \dfrac{b}{a^2+b^2} \\ -\dfrac{b}{a^2+b^2} & \dfrac{a}{a^2+b^2} \end{bmatrix}$$

Entonces

$$[D^{-1}(e^{ax} \text{ sen } bx)]_\beta = A^{-1}[e^{ax} \text{ sen } bx]_\beta = \begin{bmatrix} \dfrac{a}{a^2+b^2} & \dfrac{b}{a^2+b^2} \\ -\dfrac{b}{a^2+b^2} & \dfrac{a}{a^2+b^2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \dfrac{a}{a^2+b^2} \\ -\dfrac{b}{a^2+b^2} \end{bmatrix}$$

de donde

$$D^{-1}(e^{ax} \text{ sen } bx) = \frac{a}{a^2+b^2} e^{ax} \text{ sen } bx - \frac{b}{a^2+b^2} e^{ax} \cos bx$$

o sea

$$\int e^{ax} \text{ sen } bx\, dx = \frac{e^{ax}}{a^2+b^2} (a \text{ sen } bx - b \cos bx)$$

Por otra parte

$$[D^{-1}(e^{ax}\cos bx)]_\beta = A^{-1}[e^{ax}\cos bx]_\beta = \begin{bmatrix} \frac{a}{a^2+b^2} & \frac{b}{a^2+b^2} \\ -\frac{b}{a^2+b^2} & \frac{a}{a^2+b^2} \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{b}{a^2+b^2} \\ \frac{a}{a^2+b^2} \end{bmatrix}$$

de donde

$$D^{-1}(e^{ax}\cos bx) = \frac{b}{a^2+b^2}e^{ax}\operatorname{sen} bx + \frac{a}{a^2+b^2}e^{ax}\cos bx$$

o sea

$$\int e^{ax}\cos bx\,dx = \frac{e^{ax}}{a^2+b^2}(b\operatorname{sen} bx + a\cos bx)$$

Como último ejemplo, se va a considerar el espacio V generado por las $n + 1$ funciones $e^x, xe^x, x^2e^x, \ldots, x^ne^x$. Estas funciones son linealmente independientes y por tanto constituyen una base β de V. El objetivo es establecer una fórmula para la integral

$$\int x^ne^x\,dx$$

Las imágenes bajo D de los vectores de la base β son

$$\begin{aligned} D(e^x) &= e^x \\ D(xe^x) &= e^x + xe^x \\ D(x^2e^x) &= 2xe^x + x^2e^x \\ &\vdots \\ D(x^ne^x) &= nx^{n-1}e^x + x^ne^x \end{aligned}$$

de modo que

$$\begin{aligned} (D(e^x))_\beta &= (1, 0, 0, \ldots, 0) \\ (D(xe^x))_\beta &= (1, 1, 0, \ldots, 0) \\ (D(x^2e^x))_\beta &= (0, 2, 1, \ldots, 0) \\ &\vdots \\ (D(x^ne^x))_\beta &= (0, 0, \ldots, n, 1) \end{aligned}$$

y entonces la matriz A de orden $n + 1$ que representa a D respecto de la base β es:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 2 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & 3 & \ldots & 0 & 0 \\ & & & \ldots & & & \\ 0 & 0 & 0 & 0 & \ldots & 1 & n \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 \end{bmatrix}$$

Como el interés se encuentra solamente en la imagen inversa de la función x^ne^x, cuyo vector de coordenadas respecto de la base β es $(0, 0, \ldots, 1)$, sólo se necesita determinar la última columna de la matriz A^{-1}, la cual se denotará por Y, pues

$$[D^{-1}(x^ne^x)]_\beta = A^{-1}[x^ne^x]_\beta = Y$$

Escríbase $y_0, y_1, \ldots, y_n$ para denotar a los elementos de Y. Entonces

$$AA^{-1} = \begin{bmatrix} 1 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 2 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & 3 & \ldots & 0 & 0 \\ & & & \ldots & & & \\ 0 & 0 & 0 & 0 & \ldots & 1 & n \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 \end{bmatrix}\begin{bmatrix} & y_0 \\ & y_1 \\ * & y_2 \\ & \vdots \\ & y_{n-1} \\ & y_n \end{bmatrix} = I_{n+1}$$

de donde se obtiene el siguiente sistema de $n + 1$ ecuaciones en las $n + 1$ incógnitas $y_0, y_1, \ldots, y_n$

$$\begin{aligned} y_0 + y_1 &= 0 \\ y_1 + 2y_2 &= 0 \\ y_2 + 3y_3 &= 0 \\ &\vdots \\ y_{n-1} + ny_n &= 0 \\ y_n &= 1 \end{aligned}$$

Al comenzar con la última ecuación y sustituyendo en reversa en las restantes, se obtiene

$$\begin{aligned} y_n &= 1 \\ y_{n-1} &= -n \\ y_{n-2} &= (n)(n-1) \\ &\vdots \\ y_1 &= (-1)^{n-1}(n)(n-1)\ldots(3)(2) \\ y_0 &= (-1)^n(n)(n-1)\ldots(3)(2)(1) \end{aligned}$$

o sea que

$$y_j = (-1)^{n-j}\frac{n!}{j!} \quad j = 0, 1, \ldots, n$$

Como

$$[D^{-1}(x^ne^x)]_\beta = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}$$

se tiene que

$$D^{-1}(x^n e^x) = y_0 e^x + y_1 x e^x + y_2 x^2 e^x + \ldots + y_n x^n e^x = \sum_{j=0}^{n} y_j x^j e^x$$

$$= \sum_{j=0}^{n} (-1)^{n-j} \frac{n!}{j!} x^j e^x$$

Entonces la fórmula que se quería establecer es:

$$\int x^n e^x \, dx = \sum_{j=0}^{n} (-1)^{n-j} \frac{n!}{j!} x^j e^x$$

EJERCICIOS (SECCIÓN 6, CAPÍTULO 4)

1. En el teorema 6.1 se demostró que si T_1: $V \to U$ y T_2: $U \to W$ son transformaciones lineales inversibles, entonces $T_2 \circ T_1$: $V \to W$ es también inversible. ¿Es cierta la afirmación recíproca?

① 2. Dé un ejemplo que un operador lineal T: $V \to V$
   a) inyectivo pero no sobreyectivo.
   b) sobreyectivo pero no inyectivo.

① 3. Sean T_1 y T_2 operadores lineales en el espacio vectorial V tales que $T_1 \circ T_2$: $V \to V$ es inyectivo y sobreyectivo.
   a) Demuestre que T_2 es inyectivo.
   b) Compruebe que T_1 es sobreyectivo.
   c) ¿Es necesariamente T_2 sobreyectivo?
   d) ¿Es necesariamente T_1 inyectivo?

4. Demuestre que el operador lineal T: $\mathbf{R}^2 \to \mathbf{R}^2$ dado por

$$T(x, y) = (ax + by, cx + dy)$$

es inyectivo si, y sólo si $ad - bc \neq 0$.

① 5. Sean V y U espacios vectoriales de la misma dimensión (finita). Sea T: $V \to U$ una transformación lineal para la cual existe una transformación lineal S: $U \to V$ tal que $S \circ T = Id_V$. Pruebe que T es inversible.

6. Sean T_1: $\mathbf{R}^3 \to \mathbf{R}^2$ y T_2: $\mathbf{R}^2 \to \mathbf{R}^3$ dos transformaciones lineales. Demuestre que $T_2 \circ T_1$: $\mathbf{R}^3 \to \mathbf{R}^3$ no puede ser inversible.
   (*Sugerencia*: véase ejercicio 18 de la sección anterior. La conclusión se sigue del teorema 6.2)

7. Considere las transformaciones lineales T_1: $\mathbf{R}^n \to \mathbf{R}^m$ y T_2: $\mathbf{R}^m \to \mathbf{R}^n$. Demuestre que si $n > m$, entonces $T_2 \circ T_1$: $\mathbf{R}^n \to \mathbf{R}^n$ no puede ser inversible.

③ 8. Sean V y U espacios vectoriales de dimensión finita, suponga que T: $V \to U$ es un isomorfismo. Pruebe que la función

$$F: L(V, V) \to L(U, U) \qquad F(S) = T \circ S \circ T^{-1}$$

es un isomorfismo entre los espacios vectoriales $L(V, V)$ y $L(U, U)$. ¿Cuál es la matriz asociada a F respecto de las bases canónicas de estos espacios?

① 9. Sea V un espacio vectorial de dimensión finita. Sean T_1 y T_2 operadores lineales en V, demuestre: a) si T_1 es inversible, entonces $Ker\, T_2 = Ker(T_1 \circ T_2)$, b) si T_2 es inversible, entonces $Im\, T_1 = Im(T_1 \circ T_2)$
   [*Sugerencia*: Según el ejercicio 17 de la sección anterior solamente falta mostrar que $Ker\,(T_1 \circ T_2) \subseteq Ker\, T_2$ e $Im\, T_1 \subseteq Im(T_1 \circ T_2)$.]

10. Demuestre que el operador derivación D: $P_n \to P_n$, $D(p) = p'$ no es inversible.

① 11. Sea T un operador lineal en el espacio vectorial V tal que T^2 es el operador cero en V. Compruebe que el operador $Id - T$: $V \to V$ es no singular.

12. Demuestre que cada una de las transformaciones lineales siguientes es inversible. En cada caso, determine una fórmula explícita para la inversa de la transformación correspondiente
   a) T: $\mathbf{R}^2 \to \mathbf{R}^2$ $\quad T(x, y) = (5x - 2y, 3x + y)$
   b) T: $\mathbf{R}^3 \to \mathbf{R}^3$ $\quad T(x, y, z) = (2x + 4y - 6z, 5y - 2z, 9z)$
   c) T: $\mathbf{R}^3 \to P_2$ $\quad T(a, b, c) = a + (a + b)x + (b + c)x^2$
   d) T: $M_{2\times 2} \to M_{2\times 2}$ $\quad T(A) = A^t$
   e) T: $M_{2\times 2} \to M_{2\times 2}$ $\quad T(A) = BA$ en donde $B = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$
   f) T: $M_{2\times 2} \to P_3$ $\quad T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + b + c + d + (a + b + c)x + (a + b)x^2 + ax^3$

13. Sea V un espacio vectorial de dimensión n. Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V, considere el operador lineal T: $V \to V$ para el cual $T(v_i) = v_{n-i+1}$ $i = 1, 2, \ldots, n$. Demuestre que T es un operador inversible. Describa el operador inverso T^{-1}: $V \to V$.

① 14. Considere el operador T: $M_{2\times 2} \to M_{2\times 2}$ dado por $T(A) = AB - BA$ en donde

$$B = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$$

   a) Demuestre que el operador T no es inversible usando directamente el teorema 6.2.
   (*Sugerencia*: véase ejercicio 6 de la sección 1).
   b) Sea β la base de $M_{2\times 2}$ dada por

$$\beta = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$$

   Obtenga la matriz de T respecto de la base β. Por medio de ella, deduzca nuevamente que T no es inversible.

① 15. Sea T: $M_{n\times n} \to M_{n\times n}$ el operador lineal dado por $T(A) = BA$ en donde B es una matriz fija de orden n. Demuestre que el operador T es inversible si, y sólo si B es una matriz inversible. En tal caso, describa T^{-1}.

16. Calcule el determinante de cada uno de los siguientes operadores lineales. En cada caso, decida si el operador dado es inversible o no lo es.
   a) T: $\mathbf{R}^3 \to \mathbf{R}^3$ $\quad T(x, y, z) = (x - y + 2z, 3x + 8y - 9z, 10x - 2z)$
   b) T: $\mathbf{R}^3 \to \mathbf{R}^3$ $\quad T(x, y, z) = (x, 2x + 6y + 2z, 8x + 9y + 3z)$
   c) T: $\mathbf{R}^4 \to \mathbf{R}^4$ $\quad T(x_1, x_2, x_3, x_4) = (x_2, x_4, x_3, x_1)$
   d) T: $P_2 \to P_2$ $\quad T(a_0 + a_1 x + a_2 x^2) = (a_0 + a_1)x + a_2 x^2$
   e) T: $M_{2\times 2} \to M_{2\times 2}$ $\quad T(A) = A - A^t$
   f) T: $M_{3\times 3} \to M_{3\times 3}$ $\quad T(A) = A + A^t$

7. APLICACIONES A LA TEORÍA DE SISTEMAS DE ECUACIONES LINEALES

En esta sección se pretende usar la teoría de transformaciones lineales que se ha desarrollado en este capítulo (más concretamente, se usarán las ideas estudiadas en las secciones 2 y 6) con el fin de reobtener y precisar, por una parte, algunos de los resultados que sobre ecuaciones lineales se estudiaron en el capítulo 1, y por otra parte, también obtener nuevos resultados importantes dentro de la teoría de matrices y sistemas de ecuaciones lineales, los cuales han sido dejados para su estudio hasta este capítulo, pues es con el lenguaje de las transformaciones lineales con el que se pueden dar demostraciones simples y directas de ellos.

7.1. RANGO DE UNA MATRIZ

En la sección 7 del capítulo 3 se estudió el espacio línea de una matriz $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$. Éste fue definido como el espacio generado por los vectores línea de A.

Se van ahora a considerar los *vectores columna* de A, que se denotarán por C^A y que se definen como $C_j^A = (a_{1j}, a_{2j}, \ldots, a_{mj}) \in \mathbf{R}^m$ $j = 1, 2, \ldots, n$. Éstos son pues los vectores cuyas coordenadas coinciden con los elementos de las correspondientes columnas de A.

El *espacio columna* de A es el espacio generado por los vectores columna de A.

Obviamente se tienen las siguientes relaciones:

espacio línea de A = espacio columna de A^t
espacio línea de A^t = espacio columna de A

en donde A^t denota la transpuesta de la matriz A.

Más aún, se tiene el siguiente resultado:

TEOREMA 7.1 Sea A una matriz $m \times n$ con elementos a_{ij}. Entonces *dim* (espacio línea de A) = *dim* (espacio columna de A).

DEMOSTRACIÓN Se tiene pues la matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

Sean

$$L_i^A = (a_{i1}, a_{i2}, \ldots, a_{in}) \qquad i = 1, 2, \ldots, m$$
$$C_j^A = (a_{1j}, a_{2j}, \ldots, a_{mj}) \qquad j = 1, 2, \ldots, n$$

los correspondientes vectores línea y vectores columna de A. Supóngase que el espacio línea de A tiene dimensión k.

Sea $\beta = \{v_1, v_2, \ldots, v_k\}$ una base de tal espacio. Existen entonces escalares c_{ij} tales que

$$L_i^A = c_{i1}v_1 + c_{i2}v_2 + \ldots + c_{ik}v_k = \sum_{r=1}^{k} c_{ir}v_r \qquad i = 1, 2, \ldots, m$$

Escríbase $v_r = (b_{r1}, b_{r2}, \ldots, b_{rn})$, $r = 1, 2, \ldots, k$. Entonces

$$L_i^A = \sum_{r=1}^{k} c_{ir}v_r = \sum_{r=1}^{k} c_{ir}(b_{r1}, b_{r2}, \ldots, b_{rn})$$
$$= \left(\sum_{r=1}^{k} c_{ir}b_{r1}, \sum_{r=1}^{k} c_{ir}b_{r2}, \ldots, \sum_{r=0}^{k} c_{ir}b_{rn} \right)$$

Al igualar las j-ésimas coordenadas de los vectores de $\mathbf{R}^n$ que aparecen en ambos miembros de la expresión anterior se tiene

$$a_{ij} = \sum_{r=1}^{k} c_{ir}b_{rj} \qquad i = 1, 2, \ldots, m$$

Entonces

$$(a_{1j}, a_{2j}, \ldots, a_{mj}) = \left(\sum_{r=1}^{k} c_{1r}b_{rj}, \sum_{r=1}^{k} c_{2r}b_{rj}, \ldots, \sum_{r=1}^{k} c_{mr}b_{rj} \right)$$
$$= \sum_{r=1}^{k} b_{rj}(c_{1r}, c_{2r}, \ldots, c_{mr})$$

Es decir,

$$C_j^A = \sum_{r=1}^{k} b_{rj}u_r$$

en donde $u_r = (c_{1r}, c_{2r}, \ldots, c_{mr})$ $r = 1, 2, \ldots, k$.

Cada vector columna C_j^A se ha escrito entonces como una combinación lineal de los vectores $u_1, u_2, \ldots, u_k$. Por tanto, es claro que

$$\text{espacio columna de A} \subseteq \mathcal{L}(u_1, u_2, \ldots, u_k)$$

(véase lema previo al teorema 7.1 del capítulo 3).

Por lo tanto

$$dim \text{ (espacio columna de } A) \leq k = dim \text{ (espacio línea de } A)$$

Al repetir este mismo argumento con la matriz A^t, se concluye que

$$dim \text{ (espacio columna de } A^t) \leq dim \text{ (espacio línea de } A^t)$$

o sea

$$dim \text{ (espacio línea de } A) \leq dim \text{ (espacio columna de } A)$$

y por tanto

$$dim \text{ (espacio línea de } A) = dim \text{ (espacio columna de } A)$$

Q.E.D.

DEFINICIÓN 7.1 Sea la matriz $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$. Se llama *rango de la matriz A* a la dimensión de su espacio línea.

Según el teorema anterior, se puede definir equivalentemente el rango de una matriz como la dimensión de su espacio columna.

Obsérvese que, según el teorema 7.2 del capítulo 3, el rango de una matriz es el número de líneas no nulas en su forma escalonada reducida.

Se tiene entonces definido por una parte el concepto de rango de una matriz A. En la sección 2 de este capítulo se define el rango de una transformación lineal $T: V \to U$ como la dimensión de la imagen de ésta. Se verá a continuación que si A es la matriz que representa a $T: V \to U$ (con respecto a bases fijas de V y U), los rangos de A y de T coinciden.

TEOREMA 7.2

> Sea $T: V \to U$ una transformación lineal del espacio vectorial V de dimensión n al espacio vectorial U de dimensión m. Sea $A = [T]_{\beta_1\beta_2}$, en donde β_1 y β_2 son ciertas bases de V y U, respectivamente.
> Entonces rango de T = rango de A.

DEMOSTRACIÓN Se tiene que

$$\begin{aligned} \text{rango de } T &= dim \text{ (espacio línea de } A^t) \\ &= dim \text{ (espacio columna de } A) \\ &= dim \text{ (espacio línea de } A) \\ &= \text{rango de } A \end{aligned}$$

(la primera de estas igualdades se deduce del análisis hecho en la subsección 3.1 de este capítulo).

Q.E.D.

Se puede ahora caracterizar la inversibilidad de una matriz cuadrada en términos de su rango.

TEOREMA 7.3

> Sea A una matriz de orden n. A es inversible si, y sólo si el rango de A es n.

DEMOSTRACIÓN Considérese la transformación lineal $T: \mathbf{R}^n \to \mathbf{R}^n$ dada por $T(X) = AX$. Según el teorema 6.5, A es inversible si, y sólo si T lo es. Por otra parte, según el teorema 6.4 T es inversible si, y sólo si T es no singular, es decir, si y sólo si $Ker\ T = \{0\}$. En tal caso $dim\ Ker\ T = 0$. Según el teorema de la dimensión se tiene entonces que $dim\ Ker\ T + dim\ Im\ T = dim\ \mathbf{R}^n = n$. Por lo tanto, A es inversible, si y sólo si $dim\ Im\ T$ = rango de $T = n$ = rango de A.

Q.E.D.

El teorema anterior añade una equivalencia más de la propiedad de inversibilidad de una matriz cuadrada, a una larga lista de equivalencias de esta propiedad que han sido estudiadas desde el capítulo 1 hasta el momento. En el siguiente teorema se presenta una recopilación de estas equivalencias mencionadas anteriormente.

TEOREMA 7.4

> Sea A una matriz cuadrada de orden n. Las siguientes afirmaciones sobre la matriz A son equivalentes:
>
> 1) A es una matriz inversible.
> 2) A es equivalente por renglones a la matriz identidad I_n.
> 3) El sistema homogéneo de ecuaciones lineales $AX = 0$ tiene sólo la solución trivial.
> 4) El sistema no homogéneo de ecuaciones lineales $AX = B$ tiene una única solución para cada matriz $n \times 1$ B.
> 5) $Det\ A \neq 0$.
> 6) Los vectores línea de A son linealmente independientes.
> 7) Los vectores columna de A son linealmente independientes.
> 8) Los vectores línea de A generan a $\mathbf{R}^n$.
> 9) Los vectores columna de A generan a $\mathbf{R}^n$.
> 10) Los vectores línea de A forman una base de $\mathbf{R}^n$.
> 11) Los vectores columna de A forman una base de $\mathbf{R}^n$.
> 12) El rango de A es n.
> 13) La transformación lineal $T: V \to U$ tal que $A = [T]_{\beta\beta'}$ (β y β' bases de V y U respectivamente) es sobreyectiva.
> 14) La transformación lineal $T: V \to U$ tal que $A = [T]_{\beta\beta'}$ es inyectiva.
> 15) La transformación lineal $T: V \to U$ tal que $A = [T]_{\beta\beta'}$ es no singular.
> 16) La transformación lineal $T: V \to U$ tal que $A = [T]_{\beta\beta'}$ es inversible.

DEMOSTRACIÓN Contenida en las páginas anteriores.

Q.E.D.

Existe una serie de resultados clásicos sobre el rango de matrices que pueden ser probados fácilmente usando el lenguaje de las transformaciones lineales, al explotar el hecho de que el rango de la transformación $T(X) = AX$ es el rango de la matriz A. Algunos de estos resultados son

$$\text{rango } (A + B) \le \text{rango } A + \text{rango } B$$

$$\text{rango } AB \le \text{mín } (\text{rango } A, \text{rango } B)$$

$$\text{rango } AB \ge \text{rango } A + \text{rango } B - n$$

en donde A y B son matrices cuadradas de orden n. Se probará la tercera de ellas, dejando como ejercicio las dos restantes.

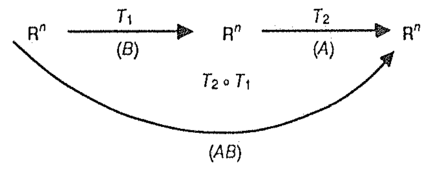
TEOREMA 7.5

> Sean A y B dos matrices cuadradas de orden n. Entonces
>
> $$\text{rango } AB \ge \text{rango } B + \text{rango } A - n$$

DEMOSTRACIÓN Considérense las transformaciones lineales T_1, T_2: $\mathbf{R}^n \to \mathbf{R}^n$ dadas por:

$$T_1(X) = BX \qquad T_2(X) = AX$$

Como se sabe, la composición $T_2 \circ T_1$: $\mathbf{R}^n \to \mathbf{R}^n$ es una transformación lineal cuya matriz asociada respecto de la base canónica de $\mathbf{R}^n$ es AB, esto es, $(T_2 \circ T_1)X = (AB)X$. Se tiene entonces una situación como la siguiente:



En el punto clave del argumento que se presenta está en considerar la transformación $\tilde{T}_2$: $ImT_1 \to \mathbf{R}^n$, definida como $\tilde{T}_2(X) = T_2(X)$. Es decir, $\tilde{T}_2$ es la restricción de la transformación lineal T_2 al subespacio ImT_1 de $\mathbf{R}^n$. Se afirma que

1) $Im(\tilde{T}_2) \subseteq Im(T_2 \circ T_1)$
2) $Ker\, \tilde{T}_2 \subseteq Ker\, T_2$

Se prueba la primera de estas relaciones. Sea $y \in Im\tilde{T}_2$. Existe entonces $z \in ImT_1$ tal que $\tilde{T}_2(z) = T_2(z) = y$. Pero si $z \in ImT_1$, existe $x \in \mathbf{R}^n$ tal que $T_1(x) = z$. Al poner todo esto junto, se ve que si $y \in Im\tilde{T}_2$, existe $x \in \mathbf{R}^n$ tal que $(T_2 \circ T_1)(x) = y$.

Entonces $y \in ImT_2 \circ T_1$ y así $Im\tilde{T}_2 \subseteq Im\tilde{T}_2 \circ T_1$.

Para probar (2), supóngase que $x \in Ker\tilde{T}_2$. Entonces $\tilde{T}_2(x) = T_2(x) = 0$. Es decir, que $x \in KerT_2$. Se tiene pues que $Ker\tilde{T}_2 \subseteq KerT_2$.

Como consecuencia de las relaciones (1) y (2) que se acaba de probar, se concluye que

$$\text{rango } (\tilde{T}_2) \le \text{rango } (T_2 \circ T_1)$$

$$\text{nulidad } (\tilde{T}_2) \le \text{nulidad } (T_2)$$

Por otra parte, el teorema de la dimensión aplicado a la transformación $\tilde{T}_2$ establece que

$$\text{dimensión del dominio de } \tilde{T}_2 = \text{rango } T_1 = \text{rango } (\tilde{T}_2) + \text{nulidad } (\tilde{T}_2)$$

Entonces

$$\begin{aligned} \text{rango } (T_2 \circ T_1) &\ge \text{rango } \tilde{T}_2 \\ &= \text{rango } T_1 - \text{nulidad } \tilde{T}_2 \\ &\ge \text{rango } T_1 - \text{nulidad } T_2 \\ &= \text{rango } T_1 - (n - \text{rango } T_2) \quad \text{(por el teorema de la dimensión aplicado a } T_2) \\ &= \text{rango } T_2 + \text{rango } T_1 - n \end{aligned}$$

Al escribir esta última relación en términos matriciales (teorema 7.2) se tiene

$$\text{rango } (AB) \ge \text{rango } B + \text{rango } A - n$$

como se deseaba.

Q.E.D.

7.2. CONDICIONES DE CONSISTENCIA DE UN SISTEMA DE ECUACIONES

En esta subsección se probará un par de teoremas relacionados con la solución de sistemas de ecuaciones lineales. La importancia de estos resultados es de carácter teórico.

Considérese primeramente el caso homogéneo.

TEOREMA 7.6

> Sea $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ una matriz $m \times n$. Considérese el sistema homogéneo de ecuaciones lineales $AX = 0$. El espacio solución de este sistema es un subespacio de $\mathbf{R}^n$ de dimensión n – rango de A. El sistema posee soluciones no triviales si, y sólo si $n >$ rango de A. Si n = rango de A, el sistema posee sólo la solución trivial.

DEMOSTRACIÓN Sea T: $\mathbf{R}^n \to \mathbf{R}^m$ la transformación lineal $T(X) = AX$. Según el teorema 3.1, se tiene: $Ker\, T$ = espacio solución de $AX = 0$, de donde nulidad de $T = dim$ (espacio solución de $AX = 0$). Por otra parte, según el teorema de la dimensión

$$\begin{aligned} \text{nulidad de } T &= dim\, R^n - \text{rango de } T \\ &= n - \text{rango de } A \end{aligned}$$

Es decir,

$$dim\ (\text{espacio solución de } AX = 0) = n - \text{rango de } A$$

como se quería probar.

Obsérvese que $n >$ rango de A = rango de T si, y sólo si nulidad de $T > 0$. En tal caso, el núcleo de T es un subespacio distinto del subespacio $\{0\}$ de $\mathbf{R}^n$; por tanto, existen vectores no nulos en *Ker T*. Tales vectores son soluciones (no triviales) del sistema $AX = 0$. Similarmente, si n = rango de A = rango de T entonces nulidad de $T = 0$. En tal caso *Ker T* = $\{0\}$ y por tanto, la única solución del sistema $AX = 0$ es la solución trivial.

Q.E.D.

Como un corolario de este teorema, reobténgase el teorema 2.1 del capítulo 1.

COROLARIO

Si el sistema de m ecuaciones con n incógnitas $AX = 0$ es tal que $m < n$, entonces el sistema posee soluciones no triviales.

DEMOSTRACIÓN Con la notación de la demostración del teorema anterior se tiene que

$$m = dim\ \mathbf{R}^m > dim\ ImT = \text{rango de } T = \text{rango de } A$$

(pues ImT es un subespacio de $\mathbf{R}^m$ —véase teorema 5.4 del capítulo 3). Entonces rango de $A \leq m < n$, y por tanto, según el teorema anterior, el sistema $AX = 0$ posee soluciones no triviales.

Q.E.D.

EJEMPLO 1 Por ejemplo, considere el sistema

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 &= 0 \\ x_1 - 3x_2 + 4x_3 - x_4 &= 0 \\ 2x_1 + 2x_2 - 2x_3 + x_4 &= 0 \\ 4x_1 + x_2 - x_3 + 2x_4 &= 0 \end{aligned}$$

Al proceder por el método de eliminación Gaussiana se obtiene

$$A = \begin{bmatrix} 1 & 2 & -3 & 2 \\ 1 & -3 & 4 & -1 \\ 2 & 2 & -2 & 1 \\ 4 & 1 & -1 & 2 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & ½ \\ 0 & 1 & 0 & -3/2 \\ 0 & 0 & 1 & -3/2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

de donde la solución del sistema es:

$$x_1 = -\frac{1}{2}t,\ x_2 = x_3 = \frac{3}{2}t,\ x_4 = t,\ t \in \mathbf{R}$$

Es decir, el vector $(x_1, x_2, x_3, x_4) \in \mathbf{R}^4$ es solución del sistema si, y sólo si

$$(x_1, x_2, x_3, x_4) = t\left(-\frac{1}{2}, \frac{3}{2}, \frac{3}{2}, 1\right)$$

de donde se ve que el espacio solución tiene dimensión 1. Obsérvese que el rango de $A = 3$ (=número de líneas no nulas en la forma escalonada reducida de A), de modo que, según el teorema 7.4 se concluye que la dimensión del espacio solución del sistema es n - rango de $A = 4 - 3 = 1$, como se ha comprobado directamente.

En el caso no homogéneo se tiene el siguiente resultado:

TEOREMA 7.7

Sea $AX = B$ un sistema de m ecuaciones lineales con n incógnitas. Este sistema es consistente si, y sólo si el rango de la matriz aumentada del sistema coincide con el rango de A.

DEMOSTRACIÓN Sea $T: \mathbf{R}^n \to \mathbf{R}^m$ la transformación lineal $T(X) = AX$. Obsérvese que el sistema $AX = B$ es consistente si, y sólo si $B \in ImT$ (esto es una consecuencia inmediata de la definición de imagen de la transformación T).

Se denotará por $[A \mid B]$ a la matriz aumentada del sistema [de orden $m \times (n + 1)$].

Supóngase entonces que el sistema $AX = B$ es consistente. Por la observación anterior se tiene que $B \in ImT$. Se afirma entonces que en este caso

$$\text{espacio columna de } A = \text{espacio columna de } [A \mid B]$$

En efecto, la contención

$$\text{espacio columna de } A \subseteq \text{espacio columna de } [A \mid B]$$

es obvia. Por otra parte, como $B \in ImT$ = espacio columna de A, se tiene también

$$\text{espacio columna de } [A \mid B] \subseteq \text{espacio columna de } A$$

lo que prueba la afirmación.

Entonces

$$\begin{aligned} \text{Rango de } A = dim\ (\text{espacio columna de } A) &= dim\ (\text{espacio columna de } [A \mid B]) \\ &= \text{rango de } [A \mid B] \end{aligned}$$

Recíprocamente, supóngase que rango de A = rango de $[A \mid B]$, es decir, las dimensiones de los espacios columna de A y columna de $[A \mid B]$ son iguales. En este caso, se debe tener necesariamente que o $B = 0$ o bien, $B \in$ espacio columna de A. En ambos casos, $B \in ImT$ y por tanto $AX = B$ es consistente.

Q.E.D.

EJERCICIOS (SECCIÓN 7, CAPÍTULO 4)

1. Determine el rango de la matriz

$$\begin{bmatrix} -3 & -1 & 2 & \lambda \\ -8 & \lambda & 5 & -1 \\ -1 & -6 & 1 & 10 \end{bmatrix}$$

en términos del parámetro λ.

2. Use el teorema 7.3 para demostrar que si $\lambda = 0$, la matriz

$$A = \begin{bmatrix} \lambda & -5 & -4 & -3 \\ 3 & 7 & 5 & 2 \\ 1 & 3 & 2 & 3 \\ 3 & 2 & 1 & -1 \end{bmatrix}$$

no es inversible.

① 3. ¿Para qué valores de λ la matriz

$$A = \begin{bmatrix} -1 & 5 & 13 & 0 \\ 2 & -6 & -16 & 1 \\ 1 & -1 & -3 & 1 \\ \lambda & 4 & 10 & 1 \end{bmatrix}$$

tiene rango mínimo?

① 4. Determine el rango de las siguientes matrices en términos del parámetro a

a) $\begin{bmatrix} 6+a & 5 & 3 & 1 \\ 2 & 4-a & 3 & 2 \\ -5 & -2 & 1-a & 1 \\ 19 & 12 & 6 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 1-2a & 1 & 1-2a \\ 1+2a & -1-a & -1+2a & 1-a \\ 2-a & 3a & 2a & 2-a \\ 5a & 1-a & 1+3a & 0 \end{bmatrix}$

① 5. Demuestre que el rango de la matriz

$$A = \begin{bmatrix} a & b & b & a \\ b & a & -a & -b \\ a+b & a+b & 2a & -2a \\ -2a & 2a & a+b & a+b \end{bmatrix}$$

es cuatro a menos que $a + b = 0$ o $b = 3a$. Encuentre el rango de A en cada uno de estos casos.

6. Encuentre los valores de t para los cuales la matriz

$$\begin{bmatrix} (1+t)t & t-1 & -t \\ 0 & 2 & -1 \\ -2t\ ) & 4-2t & t-2 \end{bmatrix}$$

tiene rango menor que 3.

② 7. Sean A y B matrices cuadradas de orden n, demuestre que

$$\text{rango}\,(A + B) \le \text{rango}\,A + \text{rango}\,B$$

8. Sean A y B matrices cuadradas de orden n, demuestre que

$$\text{rango}\,(AB) \le \text{mín}\,(\text{rango}\,A, \text{rango}\,B)$$

(*Sugerencia*: véase ejercicio 18 de la sección 5.)

② 9. Considere la transformación $T: M_{n\times n} \to M_{n\times n}$ dada por $T(A) = AB$ en donde B es una matriz dada del espacio $M_{n\times n}$. Compruebe que

$$\text{rango}\, T = n\,(\text{rango}\, B)$$

② 10. a) Sea A una matriz de orden $n \times m$ y sea B una matriz inversible de orden n. Demuestre que rango A = rango BA.
b) Si A es una matriz cuadrada de orden n y B_1 y B_2 son matrices inversibles de orden n, demuestre que rango A = rango B_1AB_2.

11. Sea A una matriz de orden n. Si la matriz B es semejante a A, demuestre que el rango de B es el mismo que el rango de A. ¿Es cierta la afirmación recíproca?

12. Para cada uno de los siguientes sistemas homogéneos de ecuaciones lineales use el teorema 7.6 para determinar la dimensión del espacio solución correspondiente

a) $\begin{aligned} 2x_1 - x_2 &= 0 \\ x_1 + x_2 &= 0 \end{aligned}$

b) $\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 2x_1 - 3x_2 + x_3 &= 0 \end{aligned}$

c) $\begin{aligned} 3x_1 + x_2 + x_3 &= 0 \\ x_1 - x_2 - x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \end{aligned}$

d) $\begin{aligned} 2x_1 - 4x_2 + 5x_3 + 3x_4 &= 0 \\ 4x_1 - 8x_2 + 17x_3 + 11x_4 &= 0 \\ 3x_1 - 6x_2 + 4x_3 + 2x_4 &= 0 \end{aligned}$

e) $\begin{aligned} 7x_1 + 12x_2 + 7x_3 &= 0 \\ 5x_1 + 8x_2 + x_3 &= 0 \\ 3x_1 + 10x_2 + 2x_3 &= 0 \\ 5x_1 + 14x_2 + 8x_3 &= 0 \end{aligned}$

f) $\begin{aligned} 4x_1 + 7x_2 + 10x_3 - 7x_4 &= 0 \\ 7x_1 + 10x_2 + 4x_3 - x_4 &= 0 \\ 7x_1 + 13x_2 + 22x_3 - 16x_4 &= 0 \\ 4x_1 + 10x_2 + 28x_3 - 22x_4 &= 0 \end{aligned}$

g) $\begin{aligned} 6x_1 + 4x_2 + 3x_3 + 5x_4 + 7x_5 &= 0 \\ 9x_1 + 6x_2 + 5x_3 + 7x_4 + 9x_5 &= 0 \\ 6x_1 + 4x_2 + 5x_3 + 3x_4 + x_5 &= 0 \\ 3x_1 + 2x_2 + \quad + 4x_4 + 8x_5 &= 0 \end{aligned}$

h) $\begin{aligned} x_1 - x_3 &= 0 \\ x_2 - x_4 &= 0 \\ -x_1 + x_3 - x_5 &= 0 \\ x_1 - x_3 + x_5 &= 0 \\ x_2 - x_4 + x_6 &= 0 \\ x_1 - x_2 + x_5 - x_6 &= 0 \end{aligned}$

13. Para cada uno de los siguientes sistemas no homogéneos de ecuaciones lineales use el teorema 7.7 para determinar si el sistema es o no consistente.

a) $$\begin{aligned} x_1 + x_2 &= 1 \\ 3x_1 - 2x_2 &= 4 \end{aligned}$$

b) $$\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ 2x_1 - x_2 + x_3 &= 3 \\ 4x_1 + x_2 + 3x_3 &= 15 \end{aligned}$$

c) $$\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ x_1 - 2x_2 + 2x_3 &= -2 \\ 3x_1 + 2x_2 - 4x_3 &= 8 \end{aligned}$$

d) $$\begin{aligned} x_1 + x_2 &= 2 \\ 3x_1 - x_2 &= 2 \\ x_1 - x_2 &= 0 \end{aligned}$$

e) $$\begin{aligned} 2x_1 + x_2 - 6x_3 &= 1 \\ x_1 + x_2 - x_3 &= 2 \\ 3x_1 + 2x_2 - 7x_3 &= 0 \end{aligned}$$

f) $$\begin{aligned} 3x_1 - 6x_2 + 3x_3 &= 9 \\ 2x_1 - 4x_2 + 2x_3 &= 6 \\ 5x_1 - 10x_2 + 5x_3 &= 15 \end{aligned}$$

g) $$\begin{aligned} 4x_1 + 3x_2 - x_3 + x_4 &= -2 \\ 2x_1 + x_2 + x_3 + x_4 &= 2 \\ 5x_1 + 3x_2 + x_3 + x_4 &= -2 \end{aligned}$$

h) $$\begin{aligned} 6x_1 + 8x_2 - 17x_3 &= 17 \\ 6x_1 + 6x_2 - 14x_3 &= 16 \\ 3x_1 + 11x_2 - 12x_3 &= 19 \\ 3x_1 + 13x_2 - 15x_3 &= 20 \end{aligned}$$

i) $$\begin{aligned} 11x_1 + 9x_2 + 7x_3 + 3x_4 &= 31 \\ 7x_1 + 5x_2 + 5x_3 + 2x_4 &= 18 \\ 7x_1 + 7x_2 + 4x_3 + 2x_4 &= 23 \\ 10x_1 + 9x_2 + 6x_3 + 3x_4 &= 33 \\ 15x_1 + 10x_2 + 10x_3 + 4x_4 &= 39 \end{aligned}$$

j) $$\begin{aligned} 2x_1 + 4x_2 + 2x_3 - 2x_4 + 2x_5 &= 1 \\ -2x_1 - 4x_2 + 2x_3 \qquad - 2x_5 &= 6 \\ 3x_1 + 6x_2 + 9x_3 - 7x_4 + 4x_5 &= 17 \\ 3x_1 + 6x_2 + 5x_3 - 5x_4 + 4x_5 &= 10 \end{aligned}$$

APÉNDICE I. FUNCIONALES LINEALES

A una transformación lineal $f: V \to \mathbf{R}$ del espacio vectorial V al campo de los reales (el cual también es un espacio vectorial-real) se le llama *funcional lineal* en V.

En este apéndice se estudiarán algunos temas relacionados con funcionales lineales. Podría parece a primera vista que los resultados que se encuentran en el estudio de los funcionales lineales son solamente particularizaciones de aquellos resultados que se estudian durante este capítulo, para transformaciones lineales en general $T: V \to U$, al caso en el que U es $\mathbf{R}$ (pues finalmente los funcionales lineales son un tipo muy es especial de transformaciones lineales). Sin embargo, se verá que existen algunos resultados muy importantes (e interesantes) relacionados con funcionales lineales en los cuales se explota precisamente la particularización del condominio de la transformación lineal.

Véase primeramente algunos ejemplos de funcionales lineales.

Si V es el espacio $\mathbf{R}^3$ y a_1, a_2 y a_3 son escalares fijos, entonces $f: \mathbf{R}^3 \to \mathbf{R}$ dado por

$$f(x, y, z) = a_1x + a_2y + a_3z$$

es un funcional lineal en $\mathbf{R}^3$.

En efecto, se comprueba fácilmente que

$$f((x, y, z) + (x', y', z')) = f(x, y, z) + f(x', y', z')$$

y

$$f(c(x, y, z)) = cf(x, y, z)$$

Más generalmente, si V es el espacio $\mathbf{R}^n$ y $a_1, a_2, \ldots, a_n$ son n escalares fijos, entonces $f: \mathbf{R}^n \to \mathbf{R}$ dado por

$$f(x_1, x_2, \ldots x_n) = a_1 x_1 + a_2x_2 + \ldots + a_nx_n$$

es un funcional lineal en $\mathbf{R}^n$.

Si $V = C([0, 1])$, el espacio de las funciones reales continuas es el intervalo [0, 1], entonces la función $f: C([0, 1]) \to \mathbf{R}$ dada por

$$f(g) = \int_0^1 g(t)dt$$

es un funcional lineal en $C([0, 1])$ pues

$$\begin{aligned} f(g_1 + g_2) &= \int_0^1 (g_1 + g_2)(t)dt = \int_0^1 (g_1(t) + g_2(t))dt \\ &= \int_0^1 g_1(t)dt + \int_0^1 g_2(t)dt \\ &= f(g_1) + f(g_2) \end{aligned}$$

y

$$f(cg) = \int_0^1 (cg)(t)dt = \int_0^1 cg(t)dt = c\int_0^1 g(t)dt = cf(g)$$

En la primera subsección de este apéndice se estudiarán "conjuntos de funcionales lineales".

AI.1. EL ESPACIO DUAL Y EL BIDUAL

En la subsección 5.1 de este capítulo se vio que el conjunto de todas las transformaciones lineales $T: V \to U$ del espacio vectorial V al espacio vectorial U, el cual se denota por $L(V, U)$, es de hecho un espacio vectorial y que si V y U eran espacios de dimensión finita su dimensión es igual al producto de las dimensiones de V y U (teorema 5.1).

En el caso particular $U = \mathbf{R}$, se tiene entonces que $L(V, \mathbf{R})$ es el espacio vectorial de todos los funcionales lineales en V. Este espacio se denota por V^* y se llama *espacio dual* de V. Las operaciones en V^* son

$$(f + g)(v) = f(v) + g(v)$$

$$(cv)(v) = cf(v), \qquad c \in \mathbf{R}, f, g \in V^*$$

Si V es de dimensión finita, V^* será también de dimensión finita y $dim\ V^* = dim\ V$, lo cual es una consecuencia del teorema 5.1 y del hecho de que $\mathbf{R}$ es un espacio vectorial (real) de dimensión 1.

También como una consecuencia del teorema 5.1 (de su demostración) se tiene que dada la base $\beta = \{v_1, v_2, \ldots, v_n\}$ del espacio vectorial V, los n funcionales lineales $f_i \in V^*$ $i = 1, 2, \ldots, n$ dados por

$$f_i(v_j) = \begin{cases} 1 & \text{si } j = i \\ 0 & \text{si } j \neq i \end{cases}$$

forman una base de V^*. Esta base es llamada *base dual de* β. Por ejemplo, si $V = \mathbf{R}^n$ y β es la base canónica de V, se tiene que, según el análisis anterior, los n funcionales $f_i : \mathbf{R}^n \to \mathbf{R}$ $i = 1, 2, \ldots, n$ dados por

$$f_i(e_j) = f_i(0, 0, \ldots, 0, \underset{j}{\underline{1}}, 0, \ldots, 0) = \begin{cases} 1 & \text{si } j = i \\ 0 & \text{si } j \neq i \end{cases}$$

constituyen la base dual de β.

Obsérvese que dado $x = (x_1, x_2, \ldots x_n) \in \mathbf{R}^n$,

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n = \sum_{j=1}^{n} x_je_j$$

se tiene que

$$f_i(x) = f_i\left(\sum_{j=1}^{n} x_je_j\right) = \sum_{j=1}^{n} x_jf_i(e_j) = x_i$$

de modo que el i-ésimo funcional f_i de la base dual de β es el funcional que a cada vector $x \in \mathbf{R}^n$ le asocia su i-ésima coordenada respecto de la base β.

Éste es un hecho general que se verifica fácilmente. En efecto, si $\beta = \{v_1, v_2, \ldots, v_n\}$ es una base del espacio vectorial V y se escribe el vector $v \in V$ como $v = c_1v_1 + c_2v_2 + \ldots + c_nv_n$, entonces si $\beta' = \{f_1, f_2, \ldots, f_n\}$ es la base dual de β se tiene que

$$f_i(v) = f_i\left(\sum_{j=1}^{n} c_jv_j\right) = \sum_{j=1}^{n} c_jf_i(v_j) = c_i = \begin{array}{l} i\text{-ésima coordenada de } v \\ \text{respecto de la base } \beta. \end{array}$$

Más aún, se tiene el siguiente resultado que dice cómo son en general los funcionales lineales pertenecientes al espacio vectorial V^* dual del espacio V de dimensión finita.

TEOREMA AI.1 Sea V un espacio de dimensión finita y sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de él. Si $f: V \to \mathbf{R}$ es un funcional lineal en V, entonces

$$f(v) = f(v_1)c_1 + f(v_2)c_2 + \ldots + f(v_n)c_n$$

en donde $(v)_\beta = (c_1, c_2, \ldots, c_n)$

DEMOSTRACIÓN Considérese la base dual $\beta' = \{f_1, f_2, \ldots, f_n\}$ de β. Dado $f \in V^*$ existen escalares $d_1, d_2, \ldots, d_n$ tales que

$$f = d_1f_1 + d_2f_2 + \ldots + d_nf_n = \sum_{i=1}^{n} d_if_i$$

Obsérvese que

$$f(v_j) = \sum_{i=1}^{n} d_if_i(v_j) = d_j$$

de modo que se puede escribir

$$f = \sum_{i=1}^{n} f(v_i)f_i$$

Si v es un vector en V y $(v)_\beta = (c_1, c_2, \ldots, c_n)$, se tiene que, por la observación previa al teorema

$$f(v) = \sum_{i=1}^{n} f(v_i)f_i(v) = \sum_{i=1}^{n} f(v_i)c_i$$

que es precisamente lo que se quería demostrar.

Q.E.D.

Dado el espacio vectorial V de dimensión finita, dígase que $\dim V = n$, se tiene entonces asociado a él su espacio dual V^*, también de dimensión n, cuyos vectores son funcionales lineales $f: V \to \mathbf{R}$. Se podría considerar similarmente el dual de este espacio vectorial V^*. Éste sería pues el dual del espacio dual de V. Este espacio es denotado como V^{**} y es llamado el bidual del espacio vectorial V.

El espacio bidual V^{**} tiene la misma dimensión que el espacio dual V^*, y por tanto, la misma dimensión que el espacio V. Los vectores en V^{**} son funcionales lineales en V^*, es decir, funcionales lineales de la forma $v: V^* \to \mathbf{R}$.

Lo único que se puede afirmar a priori sobre el espacio vectorial V y su bidual V^{**} es que éstos son espacios isomorfos, pues ambos tienen la misma dimensión. Sin embargo, se verá a continuación que existe una manera muy natural de identificar estos espacios (sin referencia alguna a bases de V y V^{**}, que es como en primera instancia se podría establecer un isomorfismo entre estos espacios) de modo que permita ver el bidual V^{**} como una "auténtica copia" de V.

A cada vector $v \in V$ se le puede asociar un funcional lineal en V^* (que también se denotará por v, el cual es entonces un elemento de V^{**}) de la siguiente manera:

$$v: V^* \to \mathbf{R}$$

$$f \to v(f) = f(v)$$

Obsérvese que para $f_1, f_2 \in V^*$ y $c \in \mathbf{R}$ se tiene

$$\begin{aligned} v(f_1 + f_2) &= (f_1 + f_2)(v) \\ &= f_1(v) + f_2(v) = v(f_1) + v(f_2) \\ v(cf_1) &= (cf_1)(v) = cf_1(v) = cv(f_1) \end{aligned}$$

de modo que efectivamente v es un funcional lineal en V^*. Es decir, el vector $v \in V$, visto como un vector del bidual V^{**}, es el funcional lineal en V^* que evaluado en $f \in V^*$ da por resultado $f(v)$. Así pues, los vectores del espacio vectorial V tienen una doble vida: 1) como elementos de V, y 2) como funcionales lineales en V^*.

Esta identificación natural es de hecho un isomorfismo.

TEOREMA AI.2

> Sea V un espacio vectorial de dimensión finita, dígase que $dim\ V = n$. Considérese la función $\varphi: V \to V^{**}$ definida como: para $v \in V$ escriba $\varphi(v) = v \in V^{**}$ en donde $v: V^* \to \mathbf{R}$, $v(f) = f(v)$. Entonces φ es un isomorfismo entre V y V^{**}.

DEMOSTRACIÓN Verifíquese primeramente que φ es lineal. Si $v_1, v_2, \in V$ y $c \in \mathbf{R}$ se tiene $\varphi(v_1 + v_2) = v_1 + v_2 \in V^{**}$ en donde $v_1 + v_2: V^* \to \mathbf{R}$ es el funcional en V^* dado por $(v_1 + v_2)f = f(v_1 + v_2)$. Como f es lineal se tiene que $(v_1 + v_2)(f) = f(v_1 + v_2) = f(v1) + f(v_2)$. Entonces $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2)$. Similarmente, $\varphi(cv_1) = cv_1 \in V^{**}$ en donde $cv_1: V^* \to \mathbf{R}$ es tal que $(cv_1)(f) = f(cv_1) = cf(v_1)$ pues f es lineal. Entonces $\varphi(cv_1) = c\, \varphi(v_1)$. Esto muestra entonces que φ es lineal.

Por otra parte el núcleo de φ es por definición

$$Ker\ \varphi = \{v \in V \mid \varphi(v) = 0 \in V^{**}\}$$

Supóngase que $\tilde{v} \in Ker\ \varphi$. Entonces $\tilde{v}: V^* \to \mathbf{R}$ es el funcional cero en V^*. Es decir, $\tilde{v}(f) = f(\tilde{v}) = 0\ \forall f \in V^*$.

Se afirma que $\tilde{v} = 0$.

En efecto, supóngase por contradicción que $\tilde{v} \neq 0$. Existen entonces vectores $u_1, u_2, \ldots, u_{n-1}, \in V$ tales que $\beta = \{\tilde{v}, u_1, \ldots, u_{n-1}\}$ es una base de V. Defínase $f \in V^*$, $f: V \to \mathbf{R}$ como $f(v) = c_1$ en donde $(v)_\beta = (c_1, c_2, \ldots, c_n)$. Claramente, $f \neq 0$ (en V^*) pues $f(\tilde{v}) = 1$. Esto contradice entonces que $v(f) = f(v) = 0$ *para todo* $f \in V^*$.

Entonces, $Ker\ \varphi = \{0\}$. Como $dim\ V = dim\ V^{**} = n$, el teorema 6.4 dice que φ es un isomorfismo.

Q.E.D.

En base a esta identificación establecida en el teorema anterior, se tiene entonces que los espacios vectoriales V y V^* juegan una relación perfectamente simétrica uno con respecto al otro: V^* es el dual de V y $V(\cong V^{**})$ es el dual de V^*. Cada vector $f \in V^*$ es un funcional lineal en V y también cada vector $v \in V$ es

(identificado naturalmente con) un funcional lineal en V^*. Esta "dualidad" es fácil de recordar con la fórmula: $v(f) = f(v)$ para $v \in V\ (=V^{**})$ y $f \in V^*$.

AI.2 LA TRANSPUESTA DE UNA TRANSFORMACIÓN LINEAL

Sean V y U espacios vectoriales de dimensión finita, dígase que $n = dim\ V$ y $m = dim\ U$. Sea $T: V \to U$ una transformación lineal. Considerando ahora los espacios duales V^* y U^*, se puede, por medio de la transformación T, definir una nueva transformación $T^t: U^* \to V^*$ de la siguiente manera: para $g \in U^*$, $T^t(g) \in V^*$ es el funcional lineal en V, $T^t(g): V \to \mathbf{R}$ dado por $(T^t(g))(v) = g(T(v))$. Esquemáticamente,

$$\begin{aligned} &\overset{T}{V \to U} \\ &v \mapsto T(v) \end{aligned}$$

$$\begin{aligned} &\overset{T^t}{U^* \to V^*} \\ &g \mapsto T^t(g) \end{aligned} \qquad \text{en donde} \qquad \begin{aligned} &T^t(g): V \to \mathbf{R} \\ &v \mapsto (T^t(g))(v) = g(T(v)) \end{aligned}$$

Es fácil verificar que T^t es una transformación lineal: si $g_1, g_2, \in U^*$ y $c \in \mathbf{R}$ se tiene para cualquier $v \in V$

$$\begin{aligned} (T^t(g_1 + g_2))(v) &= (g_1 + g_2)(T(v)) \\ &= g_1(T(v)) + g_2(T(v)) \\ &= T^t(g_1)(v) + T^t(g_2)(v) \\ &= (T^t(g_1) + T^t(g_2))(v) \end{aligned}$$

de donde

$$T^t(g_1 + g_2) = T^t(g_1) + T^t(g_2)$$

Similarmente,

$$\begin{aligned} (T^t(cg_1))(v) &= (cg_1)(T(v)) = cg_1(T(v)) \\ &= c(T^t(g_1)(v)) \\ &= (cT^t(g_1)(v)) \end{aligned}$$

de donde

$$T^t(cg_1) = c\, T^t(g_1)$$

A la transformación lineal T^t descrita anteriormente se le llama *transpuesta de la transformación T*.

Sean $\beta_1 = \{v_1, v_2, \ldots, v_n\}$ y $\beta_2 = \{u_1, u_2, \ldots u_m\}$ bases de V y U respectivamente, y considérense las correspondientes bases duales $\beta_1' = \{f_1, f_2, \ldots, f_n\}$ de V^* y $\beta_2' = \{g_1, g_2, \ldots, g_m\}$ de U^*.

Sea $A = (a_{ij})_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}$ la matriz que representa a T respecto de las bases β_1 y β_2 (esto es, $A = [T]_{\beta_1\beta_2}$) y $B = (b_{ij})_{\substack{i=1,\ldots,n\\ j=1,\ldots,m}}$ la matriz que representa a T^t respecto a las bases β_2' y β_1' (esto es, $B = [T^t]_{\beta_2'\beta_1'}$).

Se tiene entonces que

$$T(v_j) = \sum_{i=1}^{m} a_{ij}u_i$$

$$T^t(g_j) = \sum_{i=1}^{n} b_{ij}f_i$$

También, por la definición de T^t se tiene que

$$(T^t(g_j))(v_i) = g_j(T(v_i)) = g_j\left(\sum_{k=1}^{m} a_{ki}u_k\right) = \sum_{k=1}^{m} a_{ki}g_j(u_k) = a_{ji}$$

(en donde para establecer la última igualdad, se usó la definición de la base dual de β_2).

Se sabe que para cualquier funcional f en V se tiene

$$f = \sum_{i=1}^{n} f(v_i)f_i$$

(véase demostración del teorema AI.1). Entonces, para el funcional $T^t(g_j)$ en V debe tenerse

$$T^t(g_j) = \sum_{i=1}^{n} T^t(g_j)(v_i)f_i$$

Pero según lo visto anteriormente, $(T^t(g_j))(v_i) = a_{ji}$. Entonces

$$T^t(g_j) = \sum_{i=1}^{n} a_{ji}f_i$$

Comparando con

$$T^t(g_j) = \sum_{i=1}^{n} b_{ij}f_i$$

y en vista de la unicidad de la representación del vector $T^t(g_j)$ como combinación lineal de los vectores de la base β_1', se tiene que

$$a_{ji} = b_{ij}$$

y por lo tanto se concluye que $B = A^t$, o sea $[T^t]_{\beta_2'\beta_1'} = [T]^t_{\beta_1\beta_2}$. Es decir, la matriz de la transformación transpuesta de T es precisamente la transpuesta de la matriz de la transformación T.

Dada la transformación lineal $T: V \to U$, se ha visto cómo asociar a ésta su transformación transpuesta. Se podría preguntar también por la transpuesta de esta última transformación. Ésta sería entonces la transformación transpuesta de la transformación transpuesta de la transformación T, $(T^t)^t: V^{**} \to U^{**}$, definida en el bidual de V y con codominio el bidual de U. Si solamente se atiende a las matrices que representan a las transformaciones lineales correspondientes, se podría concluir que la transformación $(T^t)^t$ tiene asociada la *misma* matriz de T (pues la transpuesta de la transpuesta de una matriz es la matriz original). Lo anterior hace pensar que la transformación $(T^t)^t$ es "la misma" que la transformación T. Efectivamente, esto es cierto. En realidad este hecho no es más que el reflejo de la identificación natural de V^{**} con V y de U^{**} con U que se analizó en la subsección anterior en este apéndice, de modo entonces que la transformación $(T^t)^t : V^{**} \to U^{**}$ es una auténtica copia (en el sentido en el que V^{**} lo es de V y U^{**} lo es de U) de la transformación original $T : V \to U$.

AI.3. HIPERESPACIOS E HIPERPLANOS

Se verá aquí cómo establecer un concepto geométrico que generaliza la idea de "plano en el espacio tridimensional", usando para ello funcionales lineales.

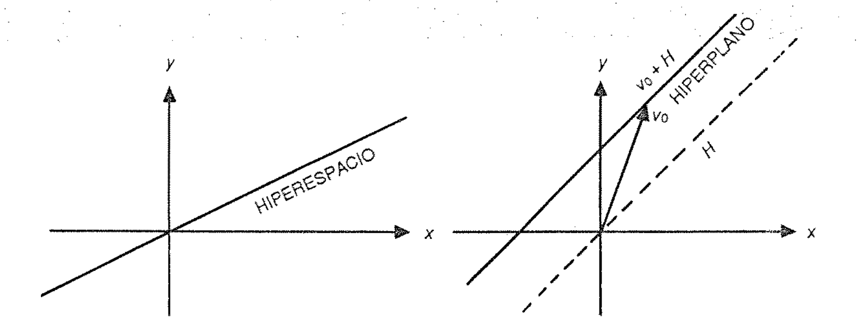
Sea $f: V \to \mathbf{R}$ un *funcional lineal no nulo* en el espacio vectorial V de dimensión n. La imagen de f es entonces un subespacio no trivial de $\mathbf{R}$. Como $0 < dim\ Im\ f \le dim\ \mathbf{R} = 1$, se tiene que $dim\ Im\ f = 1$ y entonces $Im\ f = \mathbf{R}$. Según el teorema de la dimensión se tiene

$$\text{nulidad de } f = n - 1$$

Es decir, el núcleo del funcional f es un subespacio de V cuya dimensión es menor en una unidad a la dimensión de V.

DEFINICIÓN AI.1 Sea V un espacio vectorial de dimensión n. Se dice que el subespacio H de V es un *hiperespacio de* V si $dim\ H = n - 1$. Un *hiperplano en* V es el conjunto de la forma $v_0 + H = \{v_0 + h \mid h \in H\}$ en donde H es un hiperespacio de V y v_0 es un vector en V.

Por ejemplo, en $\mathbf{R}^2$ los hiperespacios son los subespacios de dimensión $2 - 1 = 1$. Éstas son rectas que pasan por el origen. Un hiperplano es $\mathbf{R}^2$ del tipo $v_0 + H$, con H hiperespacio de $\mathbf{R}^2$, es una recta que pasa por v_0 y es paralela a la recta que representa H.



En $\mathbf{R}^3$ los hiperespacios son los subespacios de $\mathbf{R}^3$ de dimensión 3 - 1 = 2. Éstos son planos que pasan por el origen. Un hiperplano en $\mathbf{R}^3$ del tipo $v_0 + H$, con H un hiperespacio de $\mathbf{R}^3$, es un plano que pasa por v_0 y es paralelo al plano que representa H.

Según el análisis previo a la definición AI.1, todos los núcleos de funcionales lineales no nulos en un espacio vectorial V de dimensión n, son hiperespacios de él. Surge entonces de modo natural la interrogante: ¿todos los hiperespacios de un espacio vectorial V de dimensión finita son núcleos de funcionales lineales no nulos en V?

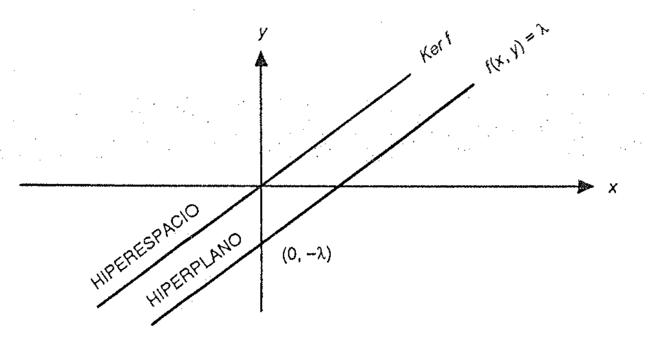
La respuesta a esta pregunta es afirmativa. Más aún, se puede usar una caracterización muy simple de los hiperplanos de un espacio vectorial por medio de funcionales lineales, de la cual se obtiene en particular la caracterización de los hiperespacios de un espacio vectorial como núcleos de funcionales lineales no nulos en V.

Antes de enunciar y probar el teorema principal de esta subsección, véanse algunos ejemplos.

Sea $V = \mathbf{R}^2$ y $f: \mathbf{R}^2 \to \mathbf{R}$ el funcional dado por $f(x, y) = 2x - y$.

El núcleo de f es por definición: $Ker\, f = \{(x, y) \mid 2x - y = 0\}$.

Éste es un hiperespacio en $\mathbf{R}^2$ que geométricamente es representado por la recta $y = 2x$. Obsérvese que al escribir $f(x, y) = \lambda$, $\lambda \in \mathbf{R}$, se obtienen los hiperplanos en $\mathbf{R}^2$, $2x - y = \lambda$, los cuales son paralelos al hiperespacio $Ker\, f$.



En $\mathbf{R}^3$ considérese el funcional lineal $f: \mathbf{R}^3 \to \mathbf{R}$ dado por

$$f(x, y, z) = x + y - z$$

El núcleo de f es el hiperespacio

$$\mathrm{Ker}\, f = \{(x, y, z) \mid x + y - z = 0\}$$

Obsérvese que al escribir $f(x, y, z) = \lambda$, $\lambda \in \mathbf{R}$, se obtienen hiperplanos paralelos al hiperespacio $Ker\, f$.

Pásese ahora al análisis que conducirá a la caracterización de los hiperplanos de un espacio vectorial por medio de funcionales lineales.

LEMA

Sea V un espacio vectorial de dimensión n y sea H un hiperespacio de él. Supóngase que $\tilde{v} \in V \backslash H$. Entonces cada vector $v \in V$ puede ser escrito en forma única como $v = \lambda\tilde{v} + h$, en donde λ es un escalar y $h \in H$.

(Obsérvese que este lema puede reenunciarse equivalentemente como: si H' es el espacio generado por $\tilde{v} \in V \backslash H$, entonces $V = H' \oplus H$.)

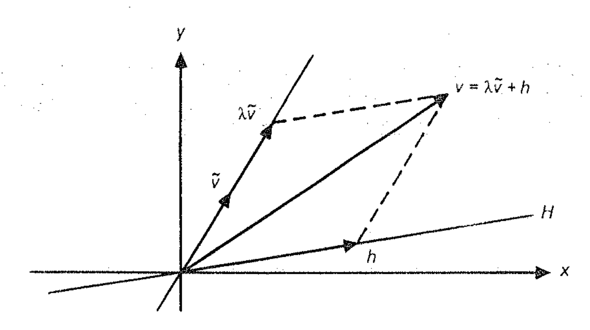
DEMOSTRACIÓN Sea $\beta_H = \{v_1, v_2, \ldots, v_{n-1}\}$ una base de H. Como $\tilde{v}$ es un vector de V que *no* pertenece a H, el conjunto $\beta = \{v_1, v_2, \ldots, v_{n-1}, \tilde{v}\}$ es una base de V, pues $dim\, V = n$ y el conjunto β es un conjunto de n vectores linealmente independientes en V (véase lema 2 sección 5 capítulo 3). Dado $v \in V$, existen únicos escalares $c_1, c_2, \ldots, c_n$ tales que

$$v = c_1v_1 + c_2v_2 + \ldots + c_{n-1}v_{n-1} + c_n\tilde{v}$$

Escríbase $h = c_1v_1 + \ldots + c_{n-1}v_{n-1}$. Es claro que $h \in H$. Poniendo $c_n = \lambda$, se obtiene la conclusión deseada del lema.

Q.E.D.

En el caso de $\mathbf{R}^2$ el contenido de este lema puede verse geométricamente según se muestra en la siguiente figura:



El siguiente teorema asegura que todos los hiperespacios en un espacio vectorial V de dimensión finita son núcleos de funcionales lineales no nulos en ese espacio.

TEOREMA AI.3

> Sea H un hiperespacio de un espacio vectorial V de dimensión n. Sea $\tilde{v} \in V \backslash H$ y c un escalar no nulo, entonces existe un único funcional lineal f: $V \to \mathbf{R}$ tal que
> a) $\text{Ker} f = H$
> b) $f(\tilde{v}) = c$

DEMOSTRACIÓN Defínase el funcional f: $V \to \mathbf{R}$ de la siguiente manera: dado $v \in V$, escríbase, según el lema previo a este teorema $v = \lambda\tilde{v} + h$, en donde $\lambda \in \mathbf{R}$ y $h \in H$ (ambos determinados por v). Defínase entonces

$$f(v) = \lambda c$$

Se afirma que f es el funcional lineal en V que cumple con la conclusión del teorema.

En efecto, si $v_1, v_2 \in V$ y $k \in \mathbf{R}$, se escribe $v_1 = \lambda_1\tilde{v} + h_1$, $v_2 = \lambda_2\tilde{v} + h_2$, de modo que $v_1 + v_2 = (\lambda_1 + \lambda_2)\tilde{v} + (h_1 + h_2)$. Entonces

$$\begin{aligned} f(v_1 + v_2) &= (\lambda_1 + \lambda_2)c \\ &= \lambda_1 c + \lambda_2 c = f(v_1) + f(v_2) \end{aligned}$$

Similarmente, $kv_1 = k\lambda_1\tilde{v} + kh_1$, y entonces

$$f(kv_1) = (k\lambda_1)c = k(\lambda_1 c) = k f(v_1)$$

lo que muestra que f es lineal.

Por otra parte, si $v \in H$, se tiene trivialmente la expresión $v = 0 \cdot \tilde{v} + v$. Como esta expresión para v es única, se concluye que $f(v) = 0 \cdot c = 0$ y por tanto, $v \in Ker\, f$. Esto muestra que $H \subset Ker\, f$. Como $dim\ H = dim\ Ker\, f = n - 1$, se concluye finalmente que $H = Ker\, f$.

También como $\tilde{v} = 1 \cdot \tilde{v} + 0$, se tiene $f(\tilde{v}) = 1 \cdot c = c$.

Véase por último que este funcional lineal es único. Supóngase que existe otro funcional lineal g: $V \to \mathbf{R}$, tal que $Ker\, g = H$ y $g(\tilde{v}) = c$. Entonces para cualquier $v \in V$, $v = \lambda\tilde{v} + h$ se tiene

$$g(v) = g(\lambda\,\tilde{v} + h) = \lambda g(\tilde{v}) + g(h) = \lambda c = f(v)$$

lo que muestra que $g = f$.

Q.E.D.

Con este teorema se responde la pregunta que se planteaba al principio de esta subsección: el subespacio H del espacio vectorial V de dimensión finita es un hiperespacio de V si, y sólo si H es el núcleo de un *funcional lineal* no nulo en V.

El siguiente teorema caracteriza a los hiperplanos de un espacio vectorial V en términos funcionales lineales en V.

TEOREMA AI.4

> Un subconjunto M del espacio vectorial V de dimensión finita es un hiperplano de V si, y sólo si
>
> $$M = \{v \in V \mid f(v) = c\}$$
>
> para algún $c \in \mathbf{R}$ y algún funcional lineal no nulo f: $V \to \mathbf{R}$.

DEMOSTRACIÓN Sea $M = \{v \in V \mid f(v) = c\}$ en donde $c \in \mathbf{R}$ y f: $V \to \mathbf{R}$ es un funcional lineal en V no nulo. Como f es no nulo, se tiene que $Im\, f = \mathbf{R}$, de modo que existe $v_0 \in V$ tal que $f(v_0) = c$. Entonces se afirma que $M = v_0 + Ker f$. En efecto, sea $v \in M$. Entonces $f(v) = c = f(v_0)$, de donde $f(v - v_0) = 0$ y por tanto, $v - v_0 \in Ker f$, o sea $v = v_0 + h$, $h \in Ker f$. Es decir, $M \subset v_0 + Ker f$.

Recíprocamente, si $v = v_0 + h$, en donde $h \in Ker f$, se tiene que $f(v) = f(v_0 + h) = f(v_0) = f(h) = c + 0 = c$, y por tanto $v \in M$. Es decir, que $v_0 + Ker f \subset M$.

Entonces $M = v_0 + Ker f$. Como $Ker f$ es hiperespacio de V, se concluye que M es hiperplano de V.

Recíprocamente, sea M un hiperplano de V, entonces $M = v_0 + H$, en donde H es un hiperespacio de V y $v_0 \in V$ (el caso de interés es cuando $v_0 \in V \backslash H$, pues si $v_0 \in H$, la conclusión de esta parte del teorema se reduce a aquella del teorema anterior, pues en este caso M sería un hiperespacio). Sea $c \in \mathbf{R}$ un escalar no nulo. Se sabe que existe un único funcional lineal f: $V \to \mathbf{R}$ tal que $Ker f = H$ y $f(v_0) = c$. Entonces, por un argumento similar al usado en la primera parte de la demostración de este teorema se concluye que $M = v_0 + H = \{v \in V \mid f(v) = c\}$

Q.E.D.

EJERCICIOS (APÉNDICE I, CAPÍTULO 4)

1. En el espacio vectorial $M_{n\times n}$, considere la función f: $M_{n\times n} \to \mathbf{R}$ dada por:

$$f(A) = tr\, A$$

Demuestre que f es un funcional lineal en $M_{n\times n}$.

2. En el espacio vectorial $C([-1, 1])$ considere la función φ: $C([-1, 1]) \to \mathbf{R}$ dada por

$$\varphi(f) = f(0)$$

Pruebe que φ es un funcional lineal en $C([-1, 1])$.

3. Considere los vectores $v_1 = (1, 1)$, $v_2 = (0, 3)$ de $\mathbf{R}^2$. Demuestre que existe un único funcional lineal f: $\mathbf{R}^2 \to \mathbf{R}$ tal que $f(v_1) = 1$, $f(v_2) = -1$. Describa explícitamente el funcional f.

4. ¿Cierto o falso? Dados dos vectores v_1 y v_2 en $\mathbf{R}^2$ existe un único funcional lineal f: $\mathbf{R}^2 \to \mathbf{R}$ tal que $f(v_1) = a$ y $f(v_2) = b$, es donde a y b son números reales dados.
5. En $\mathbf{R}^3$ considere los vectores $v_1 = (4, -1, 2)$, $v_2 = (3, 1, -1)$, $v_3 = (4, 2, 1)$. Sea f: $\mathbf{R}^3 \to \mathbf{R}$ el funcional lineal tal que $f(v_1) = f(v_2) = 1, f(v_3) = 3$, determine una fórmula explícita para f.
6. En $\mathbf{R}^3$ considere la base

$$\beta = \{(2, 1, -1), (3, 1, 2), (0, 0, 4)\}$$

Describa explícitamente la base dual de β.

7. En $M_{2\times 2}$ considere la base

$$\beta = \left\{\begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$$

Describa explícitamente la base dual de β.

8. Considere el funcional lineal f: $\mathbf{R}^2 \to \mathbf{R}$ dado por

$$f(x_1, x_2) = ax_1 + bx_2$$

en donde a y b son números reales fijos. Para cada uno de los operadores lineales T: $\mathbf{R}^2 \to \mathbf{R}^2$ dados a continuación, describa $T^t(f)$.

a) $T(x_1, x_2) = (x_1, x_2)$
b) $T(x_1, x_2) = (0, x_2)$
c) $T(x_1, x_2) = (x_1 + x_2, x_1 - x_2)$
d) $T(x_1, x_2) = (3x_1 + 4x_2, -2x_1 + 6x_2)$

9. Sea T: $M_{n\times n} \to M_{n\times n}$ el operador lineal $T(A) = BA$, en donde B es una matriz fija $M_{n\times n}$. En el ejercicio 1 se probó que f: $M_{n\times n}$ $\mathbf{R}$, $f(A) = tr\,A$ es un funcional lineal de $M_{n\times n}$. Describa $T^t(f)$.
10. En el espacio vectorial $\mathbf{R}^2$ considere el hiperespacio

$$H = \{(x, y) \mid x = y\}$$

Sea $\tilde{v} = (2, 3)$, escriba cada uno de los siguientes vectores v de $\mathbf{R}^2$ como $v = \lambda\tilde{v} + h$, $\lambda \in \mathbf{R}$, $h \in H$ (véase lema previo al teorema AI.3).

a) $v = (1, 1)$
b) $v = (4, 6)$
c) $v = (5, 5)$
d) $v = (-1, 2)$

11. En el espacio vectorial $\mathbf{R}^3$ considere el hiperespacio

$$H = \{(x, y, z) \mid x + y + z = 0\}$$

Sea $\tilde{v} = (-1, -1, -1)$. Escriba el vector $v = (3, 1, 2)$ como $v = \lambda\tilde{v} + h$, $\lambda \in \mathbf{R}$, $h \in H$.

12. En el espacio vectorial $\mathbf{R}^3$ considere el hiperespacio

$$H = \{(x, y, z) \mid 2x - y + z = 0\}$$

Describa explícitamente el funcional lineal f: $\mathbf{R}^3 \to \mathbf{R}$ tal que $Ker\,f = H$ y $f((1, 1, 1)) = 2$.

② 13. En el espacio vectorial $\mathbf{R}^3$ considere el subespacio

$$H = \mathcal{L}((2, 1, 3), (0, 1, 2))$$

a) Demuestre que H es un hiperespacio de $\mathbf{R}^3$
b) Describa explícitamente el funcional lineal f: $\mathbf{R}^3 \to \mathbf{R}$ tal que $Ker\,f = H$ y $f((1, 0, 0)) = 1$

② 14. En el espacio vectorial $M_{2\times 2}$ considere el subespacio

$$H = \mathcal{L}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right)$$

a) Compruebe que H es un hiperespacio de $M_{2\times 2}$.
b) Describa explícitamente el funcional lineal f: $M_{2\times 2} \to \mathbf{R}$ tal que $Ker\,f = H$ y $f\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right) = 1$

② 15. En el espacio vectorial $\mathbf{R}^4$ considere el subespacio

$$H = \mathcal{L}((1, 0, 1, 0), (0, 1, 0, 1), (2, 0, 0, 1))$$

a) Demuestre que H es un hiperespacio de $\mathbf{R}^4$
b) Considere el hiperplano M en $\mathbf{R}^4$ dado por

$$(0, 1, 1, 1) + H$$

Escriba M como

$$M = \{v \in \mathbf{R}^4 \mid f(v) = c\}$$

de donde c es un número y f es un funcional lineal no nulo en $\mathbf{R}^4$. (Véase teorema AI.4).

APÉNDICE II. ESPACIOS COCIENTE

En este apéndice se presenta, en el contexto del álgebra lineal, un concepto abstracto muy importante en matemáticas: el concepto de "estructura cociente".

En todo este apéndice V denotará un espacio vectorial (no necesariamente de dimensión finita) y W un subespacio de V.

En el espacio V se va a definir una relación entre sus elementos llamada (relación de) "congruencia módulo W".

DEFINICIÓN Sean v_1 y v_2 dos vectores del espacio vectorial V, se dice que v_1 *es congruente con* v_2 *módulo* W, lo cual se escribe como $v_1 \equiv v_2$ (*mod* W), si $v_1 - v_2 \in W$.

TEOREMA AII.1

> La relación de congruencia módulo W definida anteriormente es una relación de equivalencia.

DEMOSTRACIÓN a) Si $v \in V$, claramente $v \equiv v(mod\ W)$, pues $v - v = 0 \in W$ (W es un subespacio de V y por tanto contiene al cero).

b) Supóngase que $v_1 \equiv v_2\ (mod\ W)$. Entonces $v = v_1 - v_2 \in W$. Siendo W un subespacio de V, se tiene también que $-v \in W$ (el inverso aditivo de v). Entonces $-v = v_2 - v_1 \in W$, o sea que $v_2 \equiv v_1\ (mod\ W)$.

c) Si $v_1 \equiv v_2\ (mod\ W)$ y $v_2 \equiv v_3\ (mod\ W)$, entonces $v_1 - v_2 \in W$ y $v_2 - v_3 \in W$. Nuevamente por ser W un subespacio de V se tiene que $(v_1 - v_2) + (v_2 + v_3) = v_1 - v_3 \in W$. O sea que $v_1 \equiv v_3\ (mod\ W)$.

Q.E.D.

Al ser entonces la relación de congruencia modulo W una relación de equivalencia en el espacio vectorial V, éste queda dividido en clases de equivalencia disjuntas.

Considérese el conjunto

$$v + W = \{v + w \mid w \in W\}$$

Se afirma que este conjunto es precisamente la clase de equivalencia en la que se encuentra el vector $v \in V$. En efecto, las siguientes equivalencias son claras $v' \in v + W \Leftrightarrow v' = v + w$, algún $w \in W \Leftrightarrow v' - v = w \in W \Leftrightarrow v' \equiv v\ (mod\ W)$. Es decir, los elementos del conjunto $v + W$ son precisamente aquellos vectores $v' \in V$ que son congruentes con v módulo W. Se denotará por $v + W$* a la clase de equivalencia que contiene a v. (Se dice que v es "un representante" de esa clase.)

Al conjunto de todas las clases de equivalencia de V, se le denotará por V/W. Es decir,

$$V/W = \{v + W \mid v \in V\}$$

Considérese por ejemplo el espacio vectorial $V = \mathbf{R}^2$ y W el subespacio

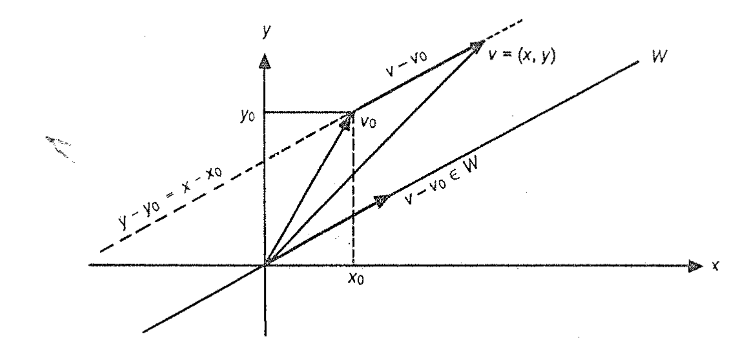
$$W = \{(x, y) \in \mathbf{R}^2 \mid x = y\}$$

(Este mismo ejemplo se usará para ilustrar otros conceptos que aparecerán en el transcurso de este apéndice.) Se describirán geométricamente a los elementos del conjunto V/W.

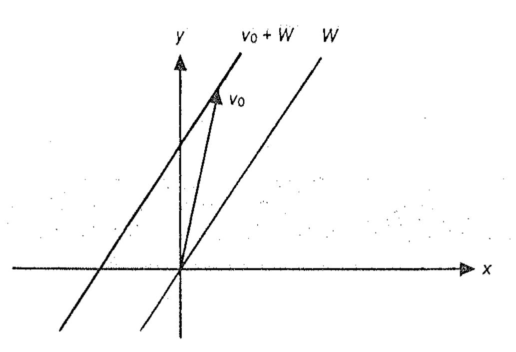
Sea $v_0 = (x_0, y_0) \in \mathbf{R}^2$, recuérdese que $v \equiv v_0(mod\ W)$ si $v - v_0 \in W$. Al escribir $v = (x, y)$ se tiene

$$v \equiv v_0(mod\ W) \Leftrightarrow (x, y) - (x_0, y_0) \in W \Leftrightarrow (x - x_0, y - y_0) \in W$$
$$\Leftrightarrow y - y_0 = x - x_0$$

Entonces, el vector $v = (x, y)$ está en la clase de equivalencia de v_0 si, y sólo si sus coordenadas satisfacen la ecuación de la recta $y - y_0 = x - x_0$, la cual es una recta que pasa por v_0 y es paralela a la recta que representa a W. Geométricamente esto se ve como



Entonces, $v_0 + W = \{(x, y) \in \mathbf{R}^2 \mid y - y_0 = x - x_0\}$. O sea,



El siguiente objetivo es mostrar que de hecho V/W es más que un simple conjunto: es un espacio vectorial.

Sean $v_1 + W, v_2 + W \in V/W$, se define la suma de $v_1 + W$ y $v_2 + W$ como

$$(v_1 + W) + (v_2 + W) = v_1 + v_2 + W$$

Es decir, la suma de la clase que contiene a v_1 y la clase que contiene a v_2, es (por definición) la clase que contiene a $v_1 + v_2$. Similarmente, si $v + W \in V/W$, y $c \in \mathbf{R}$, se define el producto de $v + W$ por el escalar $c \in \mathbf{R}$ como

$$c(v + W) = cv + W$$

Es decir, que el producto de la clase que contiene a v por el escalar $c \in \mathbf{R}$ es (por definición) la clase que contiene a cv.

Obsérvese se que se han definido estas operaciones en el conjunto V/W, usando representantes concretos de las clases de equivalencia correspondientes. Se debe entonces verificar que la definición *no depende* de las representantes que se tomen en tales clases (esto es, que se trata de una buena definición).

*Otra notación muy usada para la clase de equivalencia que contiene al vector $v \in V$ es $[v]$.

Supóngase entonces que se toma otro representante v_1' (distinto de v_1) en la clase $v_1 + W$ y otro representante v_2' (distinto de v_2) en la clase $v_2 + W$. Como $v_1 + W = v_1' + W$ y $v_2 + W = v_2' + W$, se debe verificar que $(v_1 + W) + (v_2 + W) = (v_1' + W) + (v_2' + W)$.

Como $v_1' \in v_1 + W$ y $v_2' \in v_2 + W$, se puede escribir $v_1' = v_1 + w_1$ y $v_2' = v_2 + w_2$ en donde $w_1, w_2 \in W$. Entonces

$$v_1' + v_2' = (v_1 + w_1) + (v_2 + w_2) = v_1 + v_2 + (w_1 + w_2)$$

Como $w_1 + w_2 \in W$, se tiene entonces que $v_1' + v_2' \in v_1 + v_2 + W$. Por lo tanto,

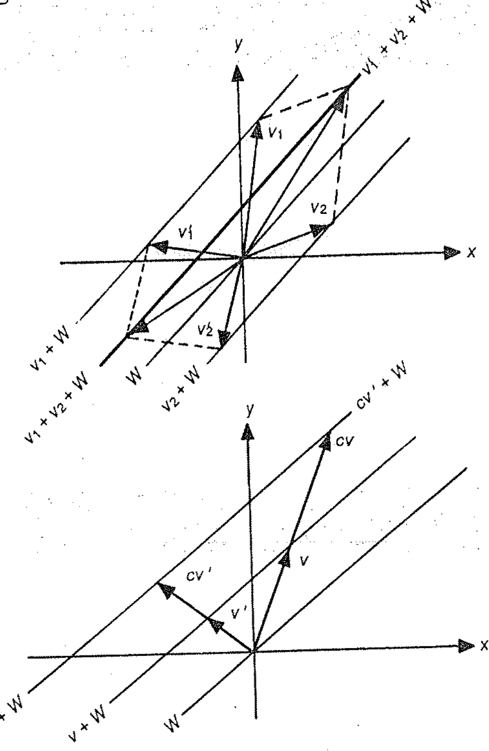
$$(v_1 + W) + (v_2 + W) = v_1 + v_2 + W = v_1' + v_2' + W = (v_1' + W) + (v_2' + W)$$

Similarmente, si $v' \in v + W$, se puede escribir $v' = v + w$, $w \in W$. Si $c \in \mathbf{R}$, se tiene $cv' = cv + cw$. Pero $cw \in W$, por lo que $cv' \in cv + W$ y entonces

$$c(v + W) = cv + W = cv' + W = c(v' + W)$$

Las operaciones de suma y producto por escalares en el conjunto V/W están entonces bien definidas.

En el ejemplo geométrico esto se ve como



TEOREMA AII.2 El conjunto V/W con las operaciones de suma y producto por escalares anteriormente definidas es un espacio vectorial.

DEMOSTRACIÓN La demostración de este teorema es un simple ejercicio de verificación de los axiomas que definen un espacio vectorial en el conjunto V/W con las operaciones que se han definido en él. Sin embargo, se darán los detalles de la demostración, pues en ellos se encuentran aspectos importantes en el manejo de las operaciones con las clases de equivalencia de V (los elementos del conjunto V/W).

Si $v_1 + W, v_2 + W, v_3 + W \in V/W$ y $\lambda, \mu \in \mathbf{R}$ se tiene

a) (la suma es conmutativa)

$$\begin{aligned}(v_1 + W) + (v_2 + W) &= v_1 + v_2 + W\\ &= v_2 + v_1 + W\\ &= (v_2 + W) + (v_1 + W)\end{aligned}$$

b) (la suma es asociativa)

$$\begin{aligned}(v_1 + W) + ((v_2 + W) + (v_3 + W)) &= (v_1 + W) + (v_2 + v_3 + W)\\ &= v_1 + (v_2 + v_3) + W\\ &= (v_1 + v_2) + v_3 + W\\ &= (v_1 + v_2 + W) + (v_3 + W)\\ &= ((v_1 + W) + (v_2 + W)) + (v_3 + W)\end{aligned}$$

c) (existe un cero en V/W). Obsérvese que la clase $0 + W$ es tal que

$$(v + W) + (0 + W) = v + 0 + W = v + W \qquad \forall\, v + W \in V/W$$

Entonces el cero de V/W es la clase de equivalencia en donde se encuentra el cero de V. Esta clase es precisamente W.

d) Dada $v + W \in V/W$, obsérvese que $-v + W \in V/W$ es tal que

$$(v + W) + (-v + W) = v - v + W = 0 + W = W \text{ (el cero de } V/W)$$

Entonces $-(v + W) = -v + W$

e) $$\begin{aligned}\lambda((v_1 + W) + (v_2 + W)) &= \lambda(v_1 + v_2 + W) = \lambda(v_1 + v_2) + W\\ &= \lambda v_1 + \lambda v_2 + W = (\lambda v_1 + W) + (\lambda v_2 + W)\\ &= \lambda(v_1 + W) + \lambda(v_2 + W)\end{aligned}$$

f) $(\lambda + \mu)(v + W) = (\lambda + \mu)\, v + W = \lambda v + \mu v + W = (\lambda v + W) + (\mu v + W)$
$= \lambda(v + W) + \mu(v + W)$

g) $(\lambda\mu)(v + W) = (\lambda\mu)v + W = \lambda(\mu v) + W = \lambda(\mu v + W) = \lambda(\mu(v + W))$

h) $1 \cdot (v + W) = 1 \cdot v + W = v + W$

Q.E.D.

Al espacio vectorial V/W se le llama *espacio cociente* ("de V entre W").

Cuando V es un espacio de dimensión finita, se tiene el siguiente resultado interesante.

TEOREMA AII.3

Sea V un espacio de dimensión n y W un subespacio de él. Entonces V/W es un espacio de dimensión finita y

$$dim\ V/W + dim\ W = dim\ V$$

DEMOSTRACIÓN Sea $\beta' = \{w_1, w_2, \ldots, w_m\}$ una base de W, existen vectores $v_1, v_2, \ldots, v_r$ en V tales que $\beta = \{w_1, w_2, \ldots, w_m, v_1, \ldots, v_r\}$ es una base de V. Se demostrará que los vectores $v_1 + W, v_2 + W, \ldots, v_r + W$ en V/W forman una base del espacio cociente V/W.

Primeramente véase que tales vectores generan a V/W. Sea $v + W$ un vector cualquiera a V/W.

Como β es una base de V, existen escalares $c_1, \ldots, c_m, d_1, \ldots, d_r$ tales que

$$v = c_1w_1 + \ldots + c_mw_m + d_1v_1 + \ldots + d_rv_r$$

de donde

$$v - (d_1v_1 + \ldots + d_rv_r) = c_1w_1 + \ldots + c_mw_m$$

El vector que aparece en el lado derecho de esta última expresión se encuentra en el subespacio W. Entonces

$$v - (d_1v_1 + \ldots + d_rv_r) \in W$$

o sea que la clase de equivalencia de v y la clase de $d_1v_1 + \ldots + d_rv_r$ es la misma. Entonces

$$\begin{aligned} v + W &= d_1v_1 + \ldots + d_rv_r + W \\ &= (d_1v_1 + W) + \ldots + (d_rv_r + W) \\ &= d_1(v_1 + W) + \ldots + d_r(v_r + W) \end{aligned}$$

lo que muestra que $v_1 + W, \ldots, v_r + W$ generan V/W.

Véase ahora que son linealmente independientes. Tómese la combinación lineal

$$d_1(v_1 + W) + \ldots + d_r(v_r + W) = 0$$

El símbolo 0 que aparece en el lado izquierdo de esta expresión se refiere a clase $0 + W$. Los vectores de esta clase son, como ya se había observado en la demostración del teorema anterior, aquellos que pertenecen a W. Entonces como

$$d_1(v_1 + W) + \ldots + d_r(v_r + W) = d_1v_1 + \ldots + d_rv_r + W$$

se tiene que

$$d_1v_1 + \ldots + d_rv_r \in W$$

Siendo β' una base de W, existen escalares $c_1, \ldots, c_m$ tales que

$$d_1v_1 + \ldots + d_rv_r = c_1w_1 + \ldots + c_mw_m$$

o lo que es lo mismo

$$c_1w_1 + \ldots + c_mw_m + (-d_1)v_1 + \ldots + (-d_r)v_r = 0$$

Pero β es una base de V. En particular, sus vectores son linealmente independientes. Entonces la expresión anterior implica que

$$c_1 = \ldots = c_m = d_1 = \ldots = d_r = 0$$

Esto muestra que los vectores $v_1 + W, \ldots, v_r + W$ son linealmente independientes y por lo tanto que ellos constituyen una base de V/W. Entonces

$$\begin{aligned} dim\ V &= r + m \\ &= dim\ V/W + dim\ W \end{aligned}$$

Q.E.D.

NOTA: La fórmula del teorema anterior tiene una fórmula particularmente sugestiva si se escribe como

$$dim\ V/W = dim\ V - dim\ W$$

Al número $dim\ V/W$ se le conoce como "*codimensión de W en V*".

Sea V un espacio vectorial y W un subespacio de él. Existe una correspondencia natural entre los elementos del espacio V y los elementos del espacio cociente V/W que le asigna a cada vector, $v \in V$ su correspondiente clase de equivalencia $v + W$

$$\begin{aligned} f: V &\to V/W \\ v &\to f(v) = v + W \end{aligned}$$

A esta función se le conoce como *proyección canónica de V en V/W*. El siguiente teorema recoge las principales propiedades de esta función.

TEOREMA AII.4

> La proyección canónica $f: V \to V/W, f(v) = v + W$, del espacio vectorial V al espacio cociente V/W, es una transformación lineal cuyo núcleo es W.

DEMOSTRACIÓN Si v_1 y v_2 son dos vectores cualesquiera de V, se tiene

$$f(v_1 + v_2) = v_1 + v_2 + W = (v_1 + W) + (v_2 + W) = f(v_1) + f(v_2)$$

Si $v \in V$ y c es un escalar

$$f(cv) = cv + W = c(v + W) = cf(v)$$

lo que muestra que f es lineal.

Por otra parte, se tiene que $Ker\, f = \{v \in V \mid f(v) = 0\}$. Es decir, $v \in Ker\, f$ si, y sólo si $f(v) = v + W$ es igual a la clase de equivalencia $0 + W$. Como ya se había observado anteriormente en el teorema AII.2, la clase $v + W$ y $0 + W$ son iguales si, y sólo si $v \in W$. Entonces $Ker\, f = W$.

Q.E.D.

El siguiente teorema es uno de los resultados más importantes en la teoría de las transformaciones lineales:

TEOREMA AII.5

> Sean V y U dos espacios vectoriales. Sea $T: V \to U$ una transformación lineal sobreyectiva de V en U. Entonces los espacios U y $V/Ker\, T$ son isomorfos. Más aún existe una correspondencia uno a uno entre el conjunto de subespacios de V que contiene a $Ker\, T$ y el conjunto de subespacios de U.

DEMOSTRACIÓN Considérese la función $\varphi: V/Ker\, T \to U$ definida de la siguiente manera: para $v + Ker\, T \in V/Ker\, T$ defínase $\varphi\,(v + Ker\, T)$ como siendo la imagen de v bajo T, o sea

$$\varphi(v + Ker\, T) = T(v)$$

Obsérvese que el vector v es sólo un representante de la clase de equivalencia $v + Ker\, T$. Se debe entonces verificar primeramente que la imagen de $v + Ker\, T$ bajo φ no depende del representante que se tome de la clase $v + Ker\, T$.

Supóngase entonces que $v' \in v + Ker\, T$. Se tiene que $v' + Ker\, T = v + Ker\, T$ y se debe mostrar entonces que $\varphi(v' + Ker\, T) = \varphi(v + Ker\, T)$.

Como $v' \in v + Ker\, T$, se tiene que $v' - v \in Ker\, T$ y por tanto, $T(v' - v) = 0$. Como T es lineal, lo anterior implica que $T(v') = T(v)$. Entonces

$$\varphi(v + Ker\, T) = T(v) = T(v') = \varphi(v' + Ker\, T)$$

lo que muestra que φ está bien definida.

Se mostrará que φ es un isomorfismo entre los espacios vectoriales $V/Ker\, T$ y U.

Primeramente véase que φ es una transformación lineal. Para $v_1 + Ker\, T$, $v_2 + Ker\, T \in V/Ker\, T$ se tiene

$$\begin{aligned}\varphi((v_1 + Ker\, T) + (v_2 + Ker\, T)) &= \varphi(v_1 + v_2 + Ker\, T)\\ &= T(v_1 + v_2) = T(v_1) + T(v_2)\\ &= \varphi(v_1 + Ker\, T) + \varphi(v_2 + Ker\, T)\end{aligned}$$

Similarmente, si $v + Ker\, T \in V/Ker\, T$ y c es un escalar

$$\varphi(c(v + Ker\, T)) = \varphi(cv + Ker\, T) = T(cv) = cT(v) = c\,\varphi(v + Ker\, T)$$

lo que muestra que φ es lineal.

Por otro lado, dado $u \in U$ existe $v \in V$ tal que $T(v) = u$, pues T es, por hipótesis, sobreyectiva. Entonces, dado $u \in U$, existe $v + Ker\, T \in V/Ker\, T$ tal que $\varphi(v + Ker\, T) = T(v) = u$, lo que muestra que φ es sobreyectiva.

Por último se ve que φ es inyectiva. Sean $v_1 + Ker\, T$, $v_2 + Ker\, T \in V/Ker\, T$ tales que $\varphi(v_1 + Ker\, T) = \varphi(v_2 + Ker\, T)$. Entonces $T(v_1) = T(v_2)$ o sea $T(v_1 - v_2) = 0$ lo que dice que $v_1 - v_2 \in Ker\, T$, o bien que $v_1 \in v_2 + Ker\, T$. Entonces $v_1 + Ker\, T = v_2 + Ker\, T$, lo que muestra que φ es inyectiva, y por tanto que φ es un isomorfismo entre $V/Ker\, T$ y U.

Pruébese ahora la segunda parte del teorema.

Considérese la función ψ definida como

$$\psi: \left\{\begin{array}{c}\text{subespacios de } V\\ \text{que contienen a}\\ Ker\, T\end{array}\right\} \to \left\{\begin{array}{c}\text{subespacios}\\ \text{de } U\end{array}\right\}$$

$$S \to \psi\,(S) = T(S)$$

Es decir, si S es un subespacio de V que contiene a $Ker\, T$, defínase $\psi\,(S)$ como $T(S) = \{u \in U \mid u = T(v), v \in S\}$. Es fácil ver que $T(S)$ es un subespacio de U. Se va a mostrar que ψ es una función biyectiva.

Sea M un subespacio de U, considérese el conjunto

$$T^{-1}(M) = \{v \in V \mid T(v) \in M\}$$

Se afirma que $T^{-1}(M)$ es un subespacio de V que contiene a $Ker\, T$ y que $\psi(T^{-1}(M)) = M$.

En efecto, si $v_1, v_2 = T^{-1}(M)$ se tiene

$$T(v_1 + v_2) = T(v_1) + T(v_2) \in M$$

pues M es un subespacio de U. Entonces $v_1 + v_2 \in T^{-1}(M)$.

Similarmente, si $v \in T^{-1}(M)$ y c es un escalar $T(cv) = cT(v) \in M$ y entonces $cv \in T^{-1}(M)$. Esto muestra que $T^{-1}(M)$ es un subespacio de U. Si $v \in Ker\ T$, entonces $T(v) = 0 \in M$ (pues M es un subespacio de U). Entonces el subespacio $T^{-1}(M)$ contiene a $Ker\ T$.

Por último se tiene que

$$\psi(T^{-1}(M)) = T(T^{-1}(M)) = M^*$$

Se ha probado entonces que ψ es sobreyectiva.

Supóngase por último que si S_1 y S_2 son dos subespacios de V que contienen a $Ker\ T$ tales que

$$\psi(S_1) = T(S_1) = T(S_2) = \psi(S_2)$$

Se afirma que $S_1 = S_2$.

Como $S_1 = T^{-1}(T(S_2))$, se tiene que $S_2 \subseteq T^{-1}(T(S_2)) = S_1$. Sea ahora $v \in T^{-1}(T(S_2))$. Entonces $T(v) \in T(S_2)$. O sea que $T(v) = T(v_0)$ para algún $v_0 \in S_2$. Entonces $v - v_0 \in Ker\ T$ (pues $T(v - v_0) = 0$). Como $Ker\ T \subseteq S_2$, se tiene $v - v_0 \in S_2$ y entonces $v \in S_2$, lo que muestra que $S_1 = T^{-1}(T(S_2)) \subseteq S_2$, quedando así probada nuestra última afirmación.

Q.E.D.

COROLARIO 1

> Sean V y U dos espacios vectoriales y $T: V \to U$ una transformación lineal. Entonces $V/Ker\ T \cong Im\ T$.

DEMOSTRACIÓN Obvio, por el teorema anterior.

Q.E.D.

COROLARIO 2

> (DE NUEVO EL TEOREMA DE LA DIMENSIÓN.) Sean V y U dos espacios vectoriales, V de dimensión finita. Sea $T: V \to U$ una transformación lineal, entonces
>
> rango de T + nulidad de $T = dim\ V$

DEMOSTRACIÓN Por el teorema AII.3 se tiene

$$dim\ V/Ker\ T - dim\ Ker\ T = dim\ V$$

Según el corolario 1, $V/Ker\ T \cong Im\ T$ y entonces $dim\ V/Ker\ T = dim\ Im\ T$ (corolario 1 del teorema 6.3). Por lo tanto

$$dim\ Im\ T + dim\ Ker\ T = dim\ V$$

*Si $f: A \to B$ es una función del conjunto A al conjunto B y si A_1 y B_1 son subconjuntos de A y B respectivamente, se tiene que: (1) $f(f^{-1}(B_1)) = B_1$, (2) $f^{-1}(f(A_1)) \supseteq A_1$.

o sea

$$\text{rango de } T + \text{nulidad de } T = dim\ V$$

Q.E.D.

Considérese por ejemplo la transformación lineal $T: \mathbf{R}^2 \to \mathbf{R}^2$ dada por

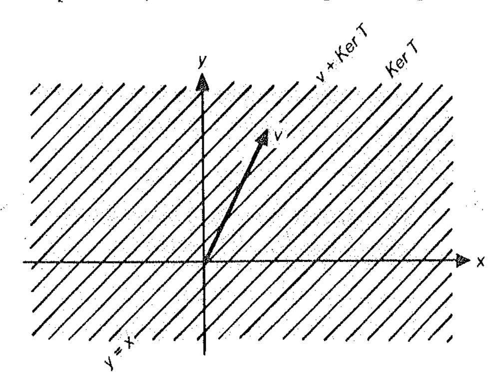
$$T(x, y) = (x - y, y - x)$$

Es fácil verificar que

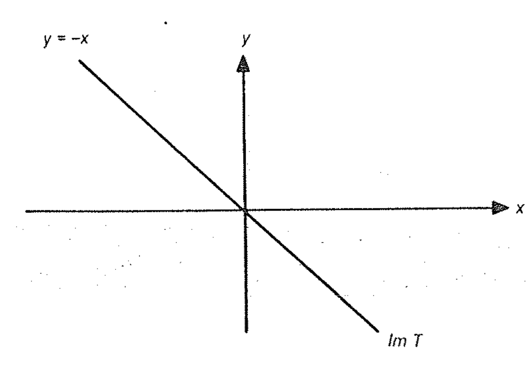
$$Ker\ T = \{(x, y) \in \mathbf{R}^2 \mid y = x\}$$

$$Im\ T = \{(x, y) \in \mathbf{R}^2 \mid y = -x\}$$

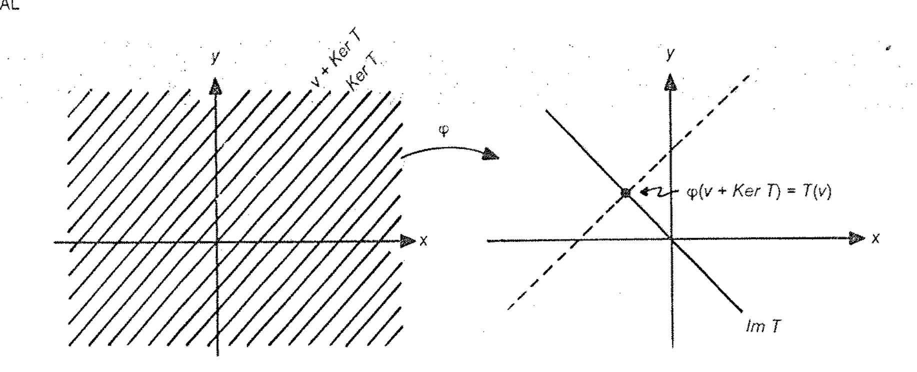
Se ha visto ya que el espacio $\mathbf{R}^2/Ker\ T$ tiene la siguiente apariencia geométrica:



Mientras que el subespacio $Im\ T$ se ve geométricamente como



El corolario 1 del teorema AII.5 asegura que $\mathbf{R}^2/Ker\ T \cong Im\ T$. No es muy difícil darse cuenta que la función φ del teorema AII.5 que establece el isomorfismo el espacio $\mathbf{R}^2/Ker\ T$ al espacio $Im\ T$ se ve en este ejemplo como



Se recomienda al lector que reconsidere la primera parte de la demostración del teorema AII.5 con este ejemplo geométrico. Más concretamente que se convenza geométricamente que la función φ: $\mathbf{R}^2/Ker\ T \to Im\ T$, $\varphi(v + Ker\ T) = T(v)$ está bien definida y es un isomorfismo.

EJERCICIOS (APÉNDICE II, CAPÍTULO 4)

1. Sea V un espacio vectorial, describa los espacios cociente V/V y $V/\{0\}$.
2. En el espacio vectorial $\mathbf{R}^3$ sea W el subespacio generado por los vectores $v_1 = (1, 0, 0)$, $v_2 = (1, 1, 0)$.
   a) ¿Cómo son los elementos del espacio cociente $\mathbf{R}^3/W$?
   b) Halle una base de $\mathbf{R}^3/W$.
3. En el espacio vectorial $\mathbf{R}^3$ sea W el subespacio generado por el vector $(0, 0, 1)$.
   a) Describa un elemento típico de $\mathbf{R}^3/W$.
   b) ¿Cómo se ve geométricamente el espacio $\mathbf{R}^3/W$?
   c) Halle una base de $\mathbf{R}^3/W$.
4. En el espacio vectorial P de todos los polinomios, considere el subespacio P_4 formado por los polinomios de grado menor o igual que 4. Describa el espacio cociente P/P_4. Halle una base de él.
5. ③ Sea V el espacio de las matrices triangulares superiores de orden 2 y U el espacio de las matrices diagonales de orden 2 (subespacio de V), demuestre que el espacio cociente V/U es isomorfo al subespacio de $M_{2\times 2}$.

$$\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| a = c = d = 0\right\}$$

6. ③ Sea V un espacio vectorial de dimensión finita y sea W un subespacio de V tal que $dim\ V/W = 1$. A los elementos del espacio cociente V/W se les llama en este caso "hiperplanos paralelos a W". Concilie esta definición con el análisis presentado en la sección 3 del apéndice I.

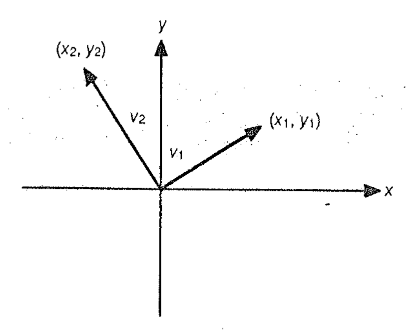
CAPÍTULO CINCO

Espacios con producto interno

Considérese el espacio vectorial $\mathbf{R}^2$.

Se ha visto que este espacio es representado geométricamente por el plano xy. En ocasiones simplemente uno se refiere al "plano $\mathbf{R}^2$", identificando al espacio con su representación geométrica.

Una característica importante en este espacio, que aún no se ha estudiado, es que en él existe una noción de "perpendicularidad" entre sus elementos. No es difícil establecer condiciones equivalentes al hecho de que los vectores $v_1 = (x_1, y_1)$ y $v_2 = (x_2, y_2)$ en $\mathbf{R}^2$ sean perpendiculares:



La pendiente de la recta en la que se encuentra v_1 es $\frac{y_1}{x_1}$. Similarmente, la pendiente de la recta en la que se encuentra v_2 es $\frac{y_2}{x_2}$. Recordando que estas rectas son perpendiculares si, y sólo si el producto de sus pendientes es igual a -1, se obtiene entonces la condición de perpendicularidad entre los vectores $v_1 = (x_1, y_1)$ y $v_2 = (x_2, y_2)$ como

$$\left(\frac{y_1}{x_1}\right)\left(\frac{y_2}{x_2}\right) = -1$$

o bien,

$$x_1x_2 + y_1y_2 = 0$$

El número $x_1x_2 + y_1y_2$ que aparece en el lado izquierdo de esta última igualdad no es más que el conocido "producto punto" de los vectores v_1 y v_2.

La función que asocia a cada par de vectores v_1 y v_2 en $\mathbf{R}^2$ su producto punto abre todo un mundo de contenido geométrico para el espacio $\mathbf{R}^2$. Escríbase esta función como

$$f: \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$$
$$(v_1, v_2) \mapsto f(v_1, v_2) = v_1 \cdot v_2$$

(en donde $v_1 \cdot v_2$ denota al producto punto de los vectores v_1 y v_2).

Obsérvese que esta función posee las siguientes propiedades:

P1) $f(v, v) \geq 0 \ \forall \ v \in \mathbf{R}^2, f(v, v) = 0$ si y sólo si $v = 0$
P2) $f(v_1, v_2) = f(v_2, v_1) \ \forall \ v_1, v_2 \in \mathbf{R}^2$
P3) $f(v_1 + v_2, v_3) = f(v_1, v_3) + f(v_2, v_3)$
P4) $f(cv_1, v_2) = cf(v_1, v_2)$

En efecto, si $v = (x, y)$ es un vector cualquiera de $\mathbf{R}^2$, se tiene

$$f(v, v) = v \cdot v = x^2 + y^2 \geq 0$$

y

$$x^2 + y^2 = 0 \Leftrightarrow x = y = 0$$

También, si $v_1 = (x_1, y_1)$ y $v_2 = (x_2, y_2)$

$$f(v_1, v_2) = v_1 \cdot v_2 = x_1x_2 + y_1y_2 = x_2x_1 + y_2y_1 = v_2 \cdot v_1 = f(v_2, v_1)$$

Por último, si $c \in \mathbf{R}$, $v_1 = (x_1, y_1)$, $v_2 = (x_2, y_2)$, $v_3 = (x_3, y_3)$

$$\begin{aligned} f(cv_1 + v_2, v_3) &= (cv_1 + v_2) \cdot v_3 = (cx_1 + x_2, cy_1 + y_2) \cdot (x_3, y_3) \\ &= (cx_1 + x_2)x_3 + (cy_1 + y_2)y_3 = cx_1x_3 + cy_1y_3 + x_2x_3 + y_2y_3 \\ &= cf(v_1, v_3) + f(v_2, v_3) \end{aligned}$$

Obsérvese también que con esta función f, se pueden *definir* los conceptos de perpendicularidad entre vectores de $\mathbf{R}^2$, magnitud de un vector en $\mathbf{R}^2$, distancia entre dos vectores de $\mathbf{R}^2$ como:

1) Los vectores $v_1, v_2 \in \mathbf{R}^2$ son perpendiculares si, y sólo si $f(v_1, v_2) = 0$.
2) La magnitud del vector $v \in \mathbf{R}$, denotada por $\|v\|$, es:

$$\|v\| = \sqrt{f(v, v)}$$

3) La distancia entre los vectores $v_1, v_2 \in \mathbf{R}^2$, denotada por $d(v_1, v_2)$ es:

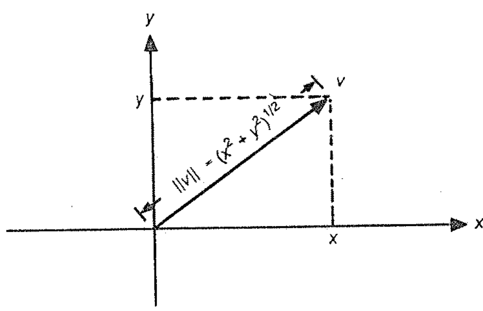
$$d(v_1, v_2) = \sqrt{f(v_1 - v_2, v_1 - v_2)}$$

En efecto, la primera de estas definiciones es clara por el análisis original.

La magnitud del vector $v = (x, y) \in \mathbf{R}^2$ según la definición (2) es:

$$\|v\| = \sqrt{f(v, v)} = \sqrt{v \cdot v} = \sqrt{x^2 + y^2}$$

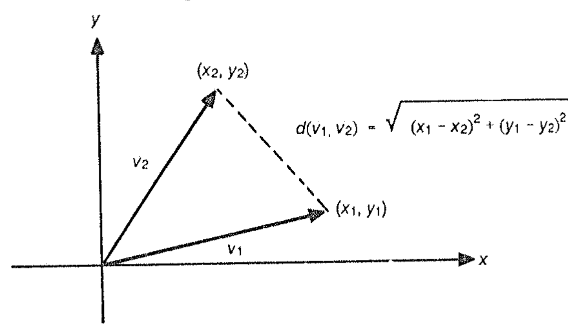
que coincide precisamente con la noción geométrica de "tamaño" del vector $v \in \mathbf{R}^2$.



Por último, si $v_1 = (x_1, y_1)$ y $v_2 = (x_2, y_2)$ se tiene, según la definición (3)

$$\begin{aligned} d(v_1, v_2) &= \sqrt{f(v_1 - v_2, v_1 - v_2)} = \sqrt{(v_1 - v_2) \cdot (v_1 - v_2)} \\ &= \sqrt{(x_1 - x_2, y_1 - y_2) \cdot (x_1 - x_2, y_1 - y_2)} \\ &= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \end{aligned}$$

que coincide también con la noción geométrica de distancia entre (x_1, y_1) y (x_2, y_2)



Todos estos conceptos (perpendicularidad entre vectores, magnitud de un vector y distancia entre vectores), los cuales pueden ser definidos por medio de la función f (el producto punto de $\mathbf{R}^2$), hacen, pues, de $\mathbf{R}^2$, un espacio vectorial con una gran riqueza geométrica.

En este capítulo uno se propone estudiar la generalización de estas ideas a espacios vectoriales abstractos.

Una función f que a cada par de vectores de un espacio vectorial le asocia un número real y que satisface las propiedades (del producto punto en $\mathbf{R}^2$) $P1$, $P2$, $P3$ y $P4$ se llama en general un "producto interno". Los espacios vectoriales para los

que existe tal función (llamados espacios con producto interno) son el objeto de estudio de este capítulo.

Se estudiarán también algunos tipos especiales de operadores lineales definidos en espacios con producto interno.

1. DEFINICIÓN Y EJEMPLOS

DEFINICIÓN 1.1 Sea V un espacio vectorial.* Un *producto interno* en V es una función $(\cdot \mid \cdot)$: $V \times V \rightarrow \mathbf{R}$ que a cada par de vectores v_1, v_2 en V le asocia el número real $(v_1 \mid v_2)$ y que satisface las siguientes propiedades:

1) $(v \mid v) \geq 0 \ \forall\, v \in V$ y $(v \mid v) = 0 \Leftrightarrow v = 0$.
2) $(v_1 \mid v_2) = (v_2 \mid v_1) \quad \forall\, v_1, v_2 \in V$.
3) $(v_1 + v_2 \mid v_3) = (v_1 \mid v_3) + (v_2 \mid v_3) \quad \forall\, v_1, v_2, v_3 \in V$.
4) $(cv_1 \mid v_2) = c(v_1 \mid v_2) \quad \forall\, c \in \mathbf{R}, \forall\, v_1, v_2 \in V$.

Un espacio vectorial V en el cual existe un producto interno $(\cdot \mid \cdot)$ es llamado *espacio con producto interno* y es denotado por $(V, (\cdot \mid \cdot))$.

En particular, si V es un espacio vectorial de dimensión finita con producto interno, se dice que V es un *espacio euclidiano*.

La propiedad (1) establece que el producto interno en V es una función *definida positiva*.

La propiedad (2) del producto interno es la propiedad de *simetría*. Se dice entonces que el producto interno es una función simétrica.

Por último, obsérvese que las propiedades (3) y (4) establecen que el producto interno es una función lineal respecto de su primera variable. Es decir, si se escribe $\varphi(\cdot) = (\cdot | v)$ en donde v es un vector fijo de V, esta función resulta ser lineal pues

$$\varphi(v_1 + v_2) = (v_1 + v_2 \mid v) = (v_1 \mid v) + (v_2 \mid v) = \varphi(v_1) + \varphi(v_2)$$

y

$$\varphi(cv_1) = (cv_1 \mid v) = c(v_1 \mid v) = c\, \varphi(v_1)$$

Más aún, debido a la propiedad de simetría del producto interno, éste también es lineal respecto de la segunda variable.

En efecto, si se escribe $\psi(\cdot) = (v \mid \cdot)$ en donde v es un vector fijo de V se tiene

$$\begin{aligned}\psi(v_1 + v_2) &= (v \mid v_1 + v_2) = (v_1 + v_2 \mid v) = (v_1 \mid v) + (v_2 \mid v) \\ &= (v \mid v_1) + (v \mid v_2) = \psi(v_1) + \psi(v_2)\end{aligned}$$

*Recuérdese que todos los espacios vectoriales considerados en este libro son reales. Cuando se consideran espacios complejos, la definición de producto interno *tiene* que modificarse. (La propiedad (2) es en este caso $(v_1|v_2) = \overline{(v_2|v_1)}$ en donde la barra denota la conjugación compleja.)

y

$$\psi(cv_1) = (v \mid cv_1) = (cv_1 \mid v) = c(v_1 \mid v) = c(v \mid v_1) = c\,\psi(v_1)$$

A una función de dos variables que es lineal respecto de cada una de ellas (es decir, es lineal como función de cada una de sus dos variables, cuando la otra se mantiene fija) se le llama *función bilineal*.

En resumen, se tiene que el producto interno es una función bilineal, simétrica y definida positiva.

En el siguiente teorema se recogen algunas consecuencias inmediatas de la definición 1.1.

TEOREMA 1.1 Sea $(V, (\cdot \mid \cdot))$ un espacio con producto interno: Entonces

1) $(0 \mid v) = 0 \quad \forall\, v \in V$.
2) $$\left(\sum_{i=1}^{n} c_i v_i \,\middle|\, \sum_{j=1}^{m} d_j u_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{m} c_i d_j (v_i \mid u_j)$$

en donde c_i, d_j son escalares y v_i, u_j, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$ son vectores de V.

DEMOSTRACIÓN Para (1) obsérvese que $(0 \mid v) = (0 \cdot w \mid v) = 0(w \mid v) = 0$ en donde w es un vector cualquiera de V (el primer símbolo 0 se refiere al vector cero de V y los siguientes tres símbolos 0 se refieren al escalar cero).

La demostración de (2) se deja como ejercicio para el lector (es una consecuencia directa de la propiedad de bilinealidad del producto interno).

Q.E.D.

Véanse algunos ejemplos.

EJEMPLO 1 En $\mathbf{R}^n$, si $x = (x_1, x_2, \ldots, x_n)$ y $y = (y_1, y_2, \ldots, y_n)$, la función

$$(x \mid y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n \tag{1.1}$$

es un producto interno (si $n = 2$, éste es el producto punto de $\mathbf{R}^2$ comentado al inicio del presente capítulo).

En efecto, se tiene

$$(x \mid x) = x_1^2 + x_2^2 + \ldots + x_n^2 \geq 0 \qquad \forall\, x \in \mathbf{R}^n$$

y

$$(x \mid x) = x_1^2 + x_2^2 + \ldots + x_n^2 = 0 \Leftrightarrow x_1 = x_2 = \ldots = x_n = 0 \Leftrightarrow x = 0$$

Por otra parte

$$(x \mid y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n = y_1x_1 + y_2x_2 + \ldots + y_nx_n = (y \mid x)$$

lo que muestra la propiedad de simetría.

También, si $x' = (x_1', x_2', \ldots, x_n') \in \mathbf{R}^n$ se tiene

$$\begin{aligned}(x + x' \mid y) &= (x_1 + x_1')y_1 + (x_2 + x_2')y_2 + \ldots + (x_n + x_n')y_n \\ &= x_1y_1 + x_2y_2 + \ldots + x_ny_n + x_1'y_1 + x_2'y_2 + \ldots + x_n'y_n \\ &= (x \mid y) + (x' \mid y)\end{aligned}$$

y por último, si $c \in \mathbf{R}$

$$\begin{aligned}(cx \mid y) &= (cx_1)y_1 + (cx_2)y_2 + \ldots + (cx_n)y_n \\ &= c(x_1y_1 + x_2y_2 + \ldots + x_ny_n) \\ &= c(x \mid y).\end{aligned}$$

Entonces la función $(\cdot \mid \cdot)$: $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ dada por (1.1) satisface las 4 propiedades de la definición 1.1 y es, por tanto, un producto interno en $\mathbf{R}^n$. Éste es llamado *producto interno canónico* (o estándar) de $\mathbf{R}^n$.

Más generalmente, si $a_1, a_2, \ldots, a_n$ son números reales positivos, la función

$$(x \mid y) = a_1x_1y_1 + a_2x_2y_2 + \ldots + a_nx_ny_n \qquad (1.2)$$

es un producto interno en $\mathbf{R}^n$. El caso $a_1 = a_2 = \ldots = a_n = 1$ corresponde al producto interno canónico de $\mathbf{R}^n$.

La verificación de que la fórmula (1.2) define de hecho un producto interno en $\mathbf{R}^n$ se deja como ejercicio para el lector (el hecho de que los números $a_1, a_2, \ldots, a_n$ sean positivos es una condición suficiente pero no necesaria para que la fórmula (1.2) defina un producto interno en $\mathbf{R}^n$, esta condición garantiza el cumplimiento de la propiedad (1) de la definición de producto interno).

No todos los productos internos de $\mathbf{R}^n$ son de la forma (1.2).

EJEMPLO 2 Por ejemplo, la fórmula

$$(x \mid y) = 3x_1y_1 - x_1y_2 - x_2y_1 + 4x_2y_2 \qquad (1.3)$$

en donde $x = (x_1, x_2)$ y $y = (y_1, y_2)$ define un producto interno en $\mathbf{R}^2$.

En efecto,

$$\begin{aligned}(x \mid x) &= 3x_1^2 - x_1x_2 - x_2x_1 + 4x_2^2 = 3x_1^2 - 2x_1x_2 + 4x_2^2 \\ &= 2x_1^2 + 3x_2^2 + (x_1 - x_2)^2 \geq 0 \qquad \forall\, x = (x_1, x_2) \in \mathbf{R}^2\end{aligned}$$

y

$$(x \mid x) = 2x_1^2 + 3x_2^2 + (x_1 - x_2)^2 = 0 \Leftrightarrow x_1 = x_2 = 0 \Leftrightarrow x = 0$$

También

$$\begin{aligned}(x \mid y) &= 3x_1y_1 - x_1y_2 - x_2y_1 + 4x_2y_2 = 3y_1x_1 - y_1x_2 - y_2x_1 + 4y_2x_2 \\ &= (y \mid x)\end{aligned}$$

Por otra parte, si $x' = (x_1', x_2')$

$$\begin{aligned}(x + x' \mid y) &= 3(x_1 + x_1')y_1 - (x_1 + x_1')y_2 - (x_2 + x_2')y_1 + 4(x_2 + x_2')y_2 \\ &= 3x_1y_1 - x_1y_2 - x_2y_1 + 4x_2y_2 + 3x_1'y_1 - x_1'y_1 - x_2'y_1 + 4x_2'y_2 \\ &= (x \mid y) + (x' \mid y)\end{aligned}$$

Por último, si $c \in \mathbf{R}$

$$\begin{aligned}(cx \mid y) &= 3(cx_1)y_1 - (cx_1)y_2 - (cx_2)y_1 + 4(cx_2)y_2 \\ &= c(3x_1y_1 - x_1y_2 - x_2y_1 + 4x_2y_2) \\ &= c(x \mid y)\end{aligned}$$

lo que muestra que efectivamente la fórmula (1.3) define un producto interno en $\mathbf{R}^2$.

Dado un espacio vectorial V de dimensión finita, éste puede heredar productos internos en $\mathbf{R}^n$ a través del isomorfismo φ: $V \to \mathbf{R}^n$, $\varphi(v) = (v)_\beta$, en donde β es una base de V, como se muestra a continuación:

Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V. Sean v y u vectores de V. Al escribir

$$v = c_1v_1 + c_2v_2 + \ldots + c_nv_n$$

$$u = d_1v_1 + d_2v_2 + \ldots + d_nv_n$$

(o sea $\varphi(v) = (v)_\beta = (c_1, c_2, \ldots, c_n)$ y $\varphi(u) = (u)_\beta = (d_1, d_2, \ldots, d_n)$), defínase la función $(\cdot \mid \cdot)_\beta$: $V \times V \to \mathbf{R}$ como

$$(v \mid u)_\beta = (\varphi(v) \mid \varphi(u)) = ((v)_\beta \mid (u)_\beta) \qquad (1.4)$$

en donde $(\cdot \mid \cdot)$ es un producto interno de $\mathbf{R}^n$.

Se afirma que esta función $(\cdot \mid \cdot)_\beta$ es un producto interno en V.

En efecto, la propiedad

$$(v \mid v)_\beta = ((v)_\beta \mid (v)_\beta) \geq 0 \qquad \forall\, v \in V$$

$$(v \mid v)_\beta = 0 \Leftrightarrow v = 0$$

es directa del hecho de que $(\cdot \mid \cdot)$ es un producto interno de $\mathbf{R}^n$.

Similarmente, la propiedad $(v \mid u)_\beta = (u \mid v)_\beta$ es también consecuencia de que $(\cdot \mid \cdot)$ es un producto interno de $\mathbf{R}^n$ pues

$$(v \mid u)_\beta = ((v)_\beta \mid (u)_\beta) = ((u)_\beta \mid (v)_\beta) = (u \mid v)_\beta$$

Para verificar las propiedades (3) y (4) de la definición (1.1) en la función $(\cdot \mid \cdot)_\beta$, se debe recordar que φ es una función lineal.

Se tiene

$$(v + v' \mid u)_\beta = (\varphi(v + v') \mid \varphi(u)) = (\varphi(v) + \varphi(v') \mid \varphi(u))$$
$$= (\varphi(v) \mid \varphi(u)) + (\varphi(v') \mid \varphi(u))$$
$$= (v \mid u)_\beta + (v' \mid u)_\beta$$

y, si $c \in \mathbf{R}$

$$(cv \mid u)_\beta = (\varphi(cv) \mid \varphi(u)) = (c\varphi(v) \mid \varphi(u)) = c(\varphi(v) \mid \varphi(u)) = c(v \mid u)_\beta$$

y queda así probado entonces que $(\cdot \mid \cdot)_\beta$ es un producto interno en V.

Por ejemplo, si $V = M_{m \times n}$ y si β es la base canónica de este espacio, se tiene que

$$(A \mid B)_\beta = \sum_{ij=1}^{m,n} a_{ij}b_{ij} \tag{1.5}$$

en donde $A = (a_{ij})_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}$, $B = (b_{ij})_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}$ es un producto interno en $M_{m \times n}$ (heredado del producto interno canónico de $\mathbf{R}^n$ por medio de la base β). Este producto interno en $M_{m \times n}$ es llamado producto interno canónico (o estándar) de $M_{m \times n}$.

Similarmente, si $V = P_n$ y β es la base canónica de este espacio, se tiene que

$$(p \mid q)_\beta = a_0b_0 + a_1b_1 + \ldots + a_nb_n \tag{1.6}$$

en donde $p = a_0 + a_1x + \ldots + a_nx^n$ y $q = b_0 + b_1x + \ldots + b_nx^n$, es un producto interno en P_n (heredado del producto interno canónico de $\mathbf{R}^n$ por medio de la base β), el cual también es llamado producto interno canónico (o estándar) de P_n.

Véase ahora un ejemplo, de un espacio vectorial de dimensión infinita con producto interno.

EJEMPLO 3 Sea $V = C([0, 1])$ el espacio de las funciones continuas en el intervalo $[0, 1]$.

Sean $f, g \in C([0, 1])$. Defínase

$$(f \mid g) = \int_0^1 f(x)g(x)dx \tag{1.7}$$

Se verá que esta fórmula define un producto interno en $C([0, 1])$.

La propiedad (2) es obvia. Las propiedades (3) y (4) se deducen a partir de propiedades de la integral definida:

$$(f_1 + f_2 \mid g) = \int_0^1 (f_1(x) + f_2(x))\, g(x)dx = \int_0^1 (f_1(x)g(x) + f_2(x)g(x))dx$$

$$= \int_0^1 f_1(x)g(x)dx + \int_0^1 f_2(x)g(x)dx = (f_1 \mid g) + (f_2 \mid g)$$

$$(cf \mid g) = \int_0^1 cf(x)g(x)dx = c\int_0^1 f(x)g(x)dx = c(f \mid g)$$

La propiedad (1) requiere un poco más de atención. Se debe probar que para cualquier $f \in C([0, 1])$ se tiene

$$(f \mid f) = \int_0^1 f(x)f(x)dx = \int_0^1 (f(x))^2dx \geq 0$$

y que

$$(f \mid f) = \int_0^1 (f(x))^2dx = 0 \Leftrightarrow f = 0$$

Si $f = 0$ (esto es, $f(x) = 0\ \forall\, x \in [0, 1]$) es claro que $(f \mid f) = 0$.

Sin embargo, la afirmación recíproca no es obvia. Se va a demostrar a continuación que si $f \neq 0$ entonces $(f \mid f) > 0$. El argumento es un argumento técnico que usa el hecho de que f es una función *continua* en $[0, 1]$.

Supóngase entonces que $f \neq 0$ (esto es, f no es el vector cero de $C([0, 1])$). Existe entonces al menos un $x_0 \in [0, 1]$ tal que $f(x_0) \neq 0$.

Debido a la continuidad de la función f, existe un $\delta > 0$ tal que $f(x) \neq 0$ para toda x en la vecindad $x_0 - \delta < x < x_0 + \delta$.* Entonces como $(f(x_0))^2 > 0$ existe un $k > 0$ tal que $(f(x))^2 > k > 0$ para $x_0 - \delta < x < x_0 + \delta$. Por lo tanto,

$$(f \mid f) = \int_0^1 (f(x))^2dx > \int_{x_0-\delta}^{x_0+\delta} (f(x))^2dx > 2k^2\delta > 0$$

como se quería demostrar.

Para finalizar esta sección, se probará una de las desigualdades más importantes en el álgebra lineal, llamada *desigualdad de Cauchy-Schwartz*.

TEOREMA 1.2 (DESIGUALDAD DE CAUCHY-SCHWARTZ.) Sea $(V, (\cdot \mid \cdot))$ un espacio con producto interno, entonces para cualesquiera vectores v y u en V se cumple la desigualdad

$$(v \mid u)^2 \leq (v \mid v)\,(u \mid u)$$

DEMOSTRACIÓN Si $v = 0$ la desigualdad es obvia. Supóngase entonces $v \neq 0$. Sea $w = tv + u \in V$, en donde $t \in \mathbf{R}$. Por la propiedad 1) del producto interno se tiene que $(w \mid w) \geq 0$,

*Si $x_0 = 0$ o 1 la vecindad se ve como $0 \leq x < \delta$ o $\delta < x \leq 1$, respectivamente. En tal caso, el argumento presentado sufre algunas modificaciones obvias.

es decir, $0 \leq (w \mid w) = (tv + u \mid tv + u)$. Al usar la propiedad de bilinealidad y de simetría del producto interno se puede escribir esta última expresión como

$$0 \leq (v \mid v)t^2 + 2(v \mid u)t + (u \mid u)$$

Esta expresión dice que la función cuadrática $f(t) = at^2 + bt + c$ (en donde $a = (v \mid v)$, $b = 2(v \mid u)$, $c = (u \mid u)$) es no negativa para todo valor de t real. Esto implica entonces que

$$b^2 - 4ac \leq 0^*$$

o sea,

$$4(v \mid u)^2 - 4(v \mid v)(u \mid u) \leq 0$$

o lo que es lo mismo

$$(v \mid u)^2 \leq (v \mid v)(u \mid u)$$

como se quería demostrar.

Q.E.D.

Cuando se particulariza el espacio $(V, (\cdot \mid \cdot))$, se obtienen algunos resultados interesantes a partir de la desigualdad de Cauchy-Schwartz.

COROLARIO 1

Sean $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ números reales arbitrarios.

Entonces $$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)$$

DEMOSTRACIÓN La conclusión del corolario es la desigualdad de Cauchy-Schwartz en $\mathbf{R}^n$ con el producto interno canónico.

Q.E.D.

*La función $f(t) = at^2 + bt + c$ representa geométricamente en el plano xy una parábola que se abre hacia arriba (pues $a = (v \mid v) > 0$). El hecho de que $f(t) \geq 0\ \forall\, t \in \mathbf{R}$ indica que la parábola *no cruza* al eje x en ningún punto. Esto implica entonces que la ecuación cuadrática $ax^2 + bx + c = 0$ tiene discriminante no positivo.

COROLARIO 2

Sean f y g dos funciones continuas en el intervalo $[0, 1]$. Entonces

$$\left(\int_0^1 f(x)g(x)dx\right)^2 \leq \left(\int_0^1 (f(x))^2dx\right)\left(\int_0^1 (g(x))^2dx\right)$$

DEMOSTRACIÓN Se trata de la desigualdad de Cauchy-Schwartz en el espacio $C([0, 1])$ con el producto interno

$$(f \mid g) = \int_0^1 f(x)g(x)dx$$

Q.E.D.

EJERCICIOS (SECCIÓN 1, CAPÍTULO 5)

1. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$, suponga que para el vector $v_0 \in V$ se tiene

$$(v_0 \mid v) = 0 \qquad \forall\, v \in V$$

Demuestre que $v_0 = 0$.

2. Sean $(\cdot \mid \cdot)_1$ y $(\cdot \mid \cdot)_2$ dos productos internos en el espacio vectorial V, demuestre que

$$(\cdot \mid \cdot) = (\cdot \mid \cdot)_1 + (\cdot \mid \cdot)_2$$

es también producto interno en V.

¿Es $(\cdot \mid \cdot) = (\cdot \mid \cdot)_1 - (\cdot \mid \cdot)_2$ también un producto interno en V?

3. Determine cuáles de las siguientes funciones $(\cdot \mid \cdot)$: $\mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ son productos internos en el espacio vectorial $\mathbf{R}^2$ ($x = (x_1, x_2)$, $y = (y_1, y_2)$)

a) $(x \mid y) = x_1y_1 - x_2y_1 - x_1y_2 + 2x_2y_2$
b) $(x \mid y) = x_1y_1 + x_2y_1 + x_1y_2 - x_2y_2$
c) $(x \mid y) = 3x_1y_1 + 10x_2y_2$
d) $(x \mid y) = x_1y_1 - x_2y_1 + x_1y_2 + 2x_2y_2$
e) $(x \mid y) = x_1^2y_1^2 + x_2^2y_2^2$

4. Determine cuáles de las siguientes funciones $(\cdot \mid \cdot)$: $\mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$ son productos internos en el espacio vectorial $\mathbf{R}^3$ ($x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$)

a) $(x \mid y) = x_1y_1 + 2x_2y_2 + 3x_3y_3$
b) $(x \mid y) = x_1y_1 + 2x_2y_2 + 3x_3y_3 + x_1y_2 + x_2y_1 + x_1y_3 + x_3y_1 + 2x_2y_3 + 2x_3y_2$
c) $(x \mid y) = x_1y_1 + x_3y_3 - x_1y_2 - x_2y_1 - x_1y_3 - x_3y_1 + x_2y_3 + x_3y_2$
d) $(x \mid y) = 2x_1y_1 + x_2y_2 - x_3y_3 + x_1y_2 + x_2y_3 + x_3y_1$
e) $(x \mid y) = x_1y_1 + 2x_2y_2 + 3x_3y_3 + 2x_1y_2 + 2x_1y_3 + 4x_2y_3$

5. Determine cuáles de las siguientes funciones $(\cdot\,|\,\cdot)$: $C[-1, 1] \times C[-1, 1] \to \mathbf{R}$ son productos internos en el espacio vectorial $C([-1, 1])$

a) $(f\,|\,g) = \int_{-1}^{1} f(x)g(x)dx$ c) $(f\,|\,g) = \int_{-1}^{1} x^2 f(x)g(x)dx$

b) $(f\,|\,g) = \int_{-1}^{1} (1 - x^2)f(x)g(x)dx$ d) $(f\,|\,g) = \int_{-1}^{1} xf(x)g(x)dx$

6. Determine cuáles de las siguientes funciones $(\cdot\,|\,\cdot)$: $M_{n\times n} \times M_{n\times n} \to \mathbf{R}$ son productos internos en el espacio vectorial $M_{n\times n}$.

a) $(A\,|\,B) = \det(AB)$

b) $(A\,|\,B) = tr\,(AB)$

c) $(A\,|\,B) = tr\,(AB^t)$

① 7. En el espacio vectorial P_n, sean

$$P_1(x) = a_0 + a_1x + \ldots + a_nx^n$$
$$P_2(x) = b_0 + b_1x + \ldots + b_nx^n$$

a) Demuestre que la función $(\cdot\,|\,\cdot)_1$: $P_n \times P_n \to \mathbf{R}$

$$(p_1\,|\,p_2)_1 = \sum_{i,j=0}^{n} \frac{a_ib_j}{i+j+1}$$

es un producto interno en P_n.

b) Compruebe que la función $(\cdot\,|\,\cdot)_2$: $P_n \times P_n \to \mathbf{R}$

$$(p_1\,|\,p_2)_2 = \int_0^1 p_1(x)p_2(x)dx$$

es un producto interno en P_n.

c) Compare las productos internos $(\cdot\,|\,\cdot)_1$ y $(\cdot\,|\,\cdot)_2$ de los dos incisos anteriores.

① 8. En el espacio vectorial P_2 considere la función $(\cdot\,|\,\cdot)$: $P_2 \times P_2 \to \mathbf{R}$ dada por

$$(p\,|\,q) = p(0)q(0) + p(1)q(1) + p(2)q(2)$$

Demuestre que esta función es un producto interno en P_2.

¿Es también un producto interno en P_1? ¿En P_3? ¿En P_n, $n \geq 4$?

9. Sea $(\cdot\,|\,\cdot)$ el producto interno canónico de $\mathbf{R}^3$, considere la base de $\mathbf{R}^3$

$$\beta = \{v_1, v_2, v_3\} = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$$

Sean c_1, c_2 y c_3 3 escalares fijos, demuestre que existe un único vector $u \in \mathbf{R}^3$ tal que

$$(u\,|\,v_i) = c_i \quad i = 1, 2, 3.$$

Determine el vector u.

10. Sea $(\cdot\,|\,\cdot)$ el producto interno canónico de $M_{2\times 2}$.

Considere la base de $M_{2\times 2}$.

$$\beta = \{v_1, v_2, v_3, v_4\} = \left\{\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ -1 & 3 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}\right\}$$

Determine la matriz $A \in M_{2\times 2}$ tal que

$$(A\,|\,v_i) = i, \qquad i = 1, 2, 3, 4.$$

11. En $\mathbf{R}^3$ considere el producto interno

$$(x\,|\,y) = x_1y_1 + 5x_2y_2 + 2x_3y_3$$

Determine al menos tres vectores $x = (x_1, x_2, x_3)$ tales que

$$(x\,|\,(1, 1, 2)) = 0$$

12. En $\mathbf{R}^3$ considere el producto interno canónico. Sean v_1 y v_2 dos vectores fijos en $\mathbf{R}^3$, demuestre que existen una infinidad de vectores $x = (x_1, x_2, x_3) \in \mathbf{R}^3$ tales que

$$(x\,|\,v_1) = (x\,|\,v_2) = 0$$

13. Sea V un espacio vectorial con el producto interno $(\cdot\,|\,\cdot)$. Sean u, v dos vectores cualesquiera de V.

Si $v \neq 0$, sea $w \in V$ el vector definido por

$$w = u - \frac{(v\,|\,u)}{(v\,|\,v)}v$$

Demuestre que

$$(w\,|\,w) = (u\,|\,u) - \frac{(v\,|\,u)}{(v\,|\,v)}(v\,|\,u)$$

Concluya de esta última expresión la desigualdad de Cauchy-Schwartz.

14. Compruebe que para cualesquiera números reales x_1, x_2, y_1, y_2 se cumple la desigualdad

$$(3x_1y_1 - x_1y_2 - x_2y_1 + 4x_2y_2)^2 \leq (3x_1^2 - 2x_1x_2 + 4x_2^2)(3y_1^2 - 2y_1y_2 + 4y_2^2)$$

(*Sugerencia*: Considere $\mathbf{R}^2$ con el producto interno introducido en la pág. 406).

15. Sean $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ números reales cualesquiera y sean $c_1, c_2, \ldots, c_n$ números reales positivos, demuestre que

$$\left(\sum_{i=1}^{n} c_ix_iy_i\right)^2 \leq \left(\sum_{i=1}^{n} c_ix_i^2\right)\left(\sum_{i=1}^{n} c_iy_i^2\right)$$

16. Pruebe que la igualdad en la desigualdad de Cauchy-Schwartz se cumple si, y sólo si los vectores son linealmente dependientes.

② 17. Sea V un espacio vectorial con el producto interno $(\cdot\,|\,\cdot)$. Sean $v_1, v_2, \ldots, v_n$ vectores de V. Se define el *determinante de Gram* de los vectores $v_1, v_2, \ldots, v_n$, denotado por $G(v_1, v_2, \ldots, v_n)$, como

$$G(v_1, v_2, \ldots, v_n) = det\begin{bmatrix} (v_1\,|\,v_1) & (v_1\,|\,v_2) & \ldots & (v_1\,|\,v_n) \\ (v_2\,|\,v_1) & (v_2\,|\,v_2) & \ldots & (v_2\,|\,v_n) \\ & & & \\ (v_n\,|\,v_1) & (v_n\,|\,v_2) & \ldots & (v_n\,|\,v_n) \end{bmatrix}$$

El objetivo de este ejercicio es demostrar que los vectores $v_1, v_2, \ldots, v_n$ son linealmente dependientes si, y sólo si $G(v_1, v_2, \ldots, v_n) = 0$.

a) Demuestre que si $n = 2$ la afirmación anterior es equivalente al ejercicio 16.

b) Considere el sistema homogéneo de ecuaciones lineales $AX = 0$, en donde $A = ((v_i|v_j))_{i,j=1,2,\ldots,n}$. Observe que $\det A = G(v_1, v_2, \ldots, v_n)$. Demuestre que este sistema se puede escribir como:

$$\left(v_1 \,\middle|\, \sum_{i=1}^{n} x_i v_i\right) = 0$$

$$\left(v_2 \,\middle|\, \sum_{i=1}^{n} x_i v_i\right) = 0$$

$$\ldots$$

$$\left(v_n \,\middle|\, \sum_{i=1}^{n} x_i v_i\right) = 0$$

en donde $x = (x_1, x_2, \ldots, x_n)$. Concluya entonces que los vectores $v_1, v_2, \ldots, v_n$ son linealmente dependientes si, y sólo si $G(v_1, v_2, \ldots, v_n) = 0$.

2. NORMA Y DISTANCIA

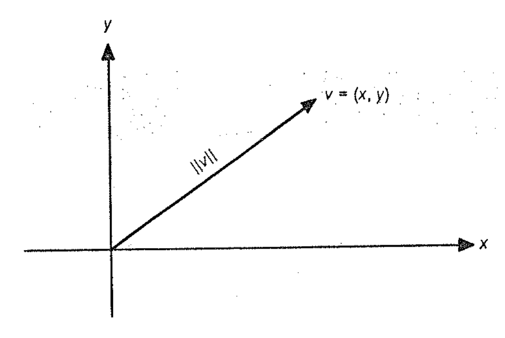
Sea $(V, (\cdot|\cdot))$ un espacio con producto interno. Se define la *norma* del vector $v \in V$, denotada por $\|v\|$, como

$$\|v\| = \sqrt{(v\,|\,v)}$$

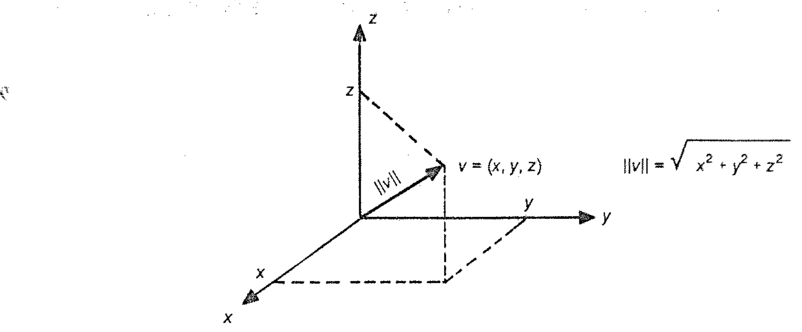
Obsérvese que esta definición tiene perfecto sentido, pues para cualquier $v \in V$ se tiene $(v\,|\,v) \geq 0$, de modo que siempre es posible extraer raíz cuadrada del número $(v\,|\,v)$.

Si $V = \mathbf{R}^2$ y $(\cdot|\cdot)$ es el producto interno canónico, la norma del vector $v \in \mathbf{R}^2$ no es más que la medida natural del tamaño de v, pues

$$\|(x, y)\| = \sqrt{((x, y)\,|\,(x, y))} = \sqrt{x^2 + y^2}$$



Un resultado similar se tiene en el espacio $\mathbf{R}^3$ con el producto interno canónico



En general, si $V = \mathbf{R}^n$ y $(\cdot|\cdot)$ es el producto interno canónico de este espacio se tiene que

$$\|v\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

en donde $v = (x_1, x_2, \ldots, x_n)$.

El concepto de norma de un vector v en un espacio con producto interno generaliza entonces la idea natural de "tamaño" de un vector en $\mathbf{R}^2$ o $\mathbf{R}^3$.

EJEMPLO 1 Por ejemplo, si $V = M_{2\times 2}$ con el producto interno canónico se tiene

$$\left\|\begin{bmatrix} 1 & 2 \\ 1 & -3 \end{bmatrix}\right\| = \sqrt{(1)^2 + (2)^2 + (1)^2 + (-3)^2} = \sqrt{15}$$

Si $V = C([0, 1])$ con el producto interno

$$(f\,|\,g) = \int_0^1 f(x)g(x)dx$$

se tiene en general, para $f \in C([0, 1])$

$$\|f\| = \sqrt{(f|f)} = \sqrt{\int_0^1 f(x)f(x)dx} = \sqrt{\int_0^1 (f(x))^2 dx}$$

de modo que, por ejemplo

$$\|e^x\| = \sqrt{\int_0^1 e^{2x}dx} = \sqrt{\frac{1}{2}(e^2 - 1)}$$

$$\|\text{sen}\, x\| = \sqrt{\int_0^1 \text{sen}^2 x dx} = \sqrt{\frac{1}{2} - \frac{1}{4}\text{sen}2}$$

Obsérvese que la norma de un vector v en el espacio con producto interno V, depende del producto interno que se tenga definido en V.

Por ejemplo, tómese el espacio $V = \mathbf{R}^2$ y el vector $v = (2, 3) \in \mathbf{R}^2$. Si en $\mathbf{R}^2$ se tiene el producto interno canónico, entonces

$$\|(2, 3)\| = \sqrt{(2)^2 + (3)^2} = \sqrt{13}$$

Sin embargo, si en $\mathbf{R}^2$ se tiene el producto interno

$$(v \mid u) = 6x_1y_1 + 8x_2y_2$$

en donde $v = (x_1, x_2)$ y $u = (y_1, y_2)$, se tiene

$$\|(2, 3)\| = \sqrt{((2, 3) \mid (2, 3))} = \sqrt{6(2)^2 + 8(3)^2} = \sqrt{96}$$

Obsérvese también que con el concepto de norma de un vector se puede reescribir la desigualdad de Cauchy-Schwartz vista en la sección anterior como

$$|(v \mid u)| \leq \|v\| \; \|u\|$$

En el siguiente teorema se presentan las propiedades más importantes de que goza la norma de un vector en un espacio vectorial con producto interno.

TEOREMA 2.1

Sea $(V, (\cdot \mid \cdot))$ un espacio vectorial con producto interno, se tienen las siguientes propiedades para la norma $\|\cdot\| = \sqrt{(\cdot \mid \cdot)}$

1) $\|v\| \geq 0 \quad \forall\, v \in V$ y $\|v\| = 0 \Leftrightarrow v = 0$
2) $\|cv\| = |c| \; \|v\| \quad \forall\, c \in \mathbf{R}, \quad \forall\, v \in V.$
3) $\|v + u\| \leq \|v\| + \|u\| \quad \forall\, v, u \in V.$

DEMOSTRACIÓN La propiedad (1) se sigue inmediatamente de la propiedad

$$(v \mid v) \geq 0 \quad \forall\, v \in V. \qquad (v \mid v) = 0 \quad \Leftrightarrow \quad v = 0$$

del producto interno.

Veamos la validez de (2)

$$\|cv\| = \sqrt{(cv \mid cv)} = \sqrt{c^2\,(v \mid v)} = \sqrt{c^2}\,\sqrt{(v \mid v)} = |c|\,\|v\|$$

Por último, se tiene

$$\|v + u\|^2 = (v + u \mid v + u) = (v \mid v) + 2(v \mid u) + (u \mid u)$$
$$= \|v\|^2 + 2(v \mid u) + \|u\|^2$$

Pero $(v \mid u) \leq |(v \mid u)| \leq \|v\|\,\|u\|$ (desigualdad de Cauchy-Schwartz). Entonces

$$\|v + u\|^2 \leq \|v\|^2 + 2\|v\| \; \|u\| + \|u\|^2 = (\|v\| + \|u\|)^2$$

o sea,

$$\|v + u\| \leq \|v\| + \|u\|$$

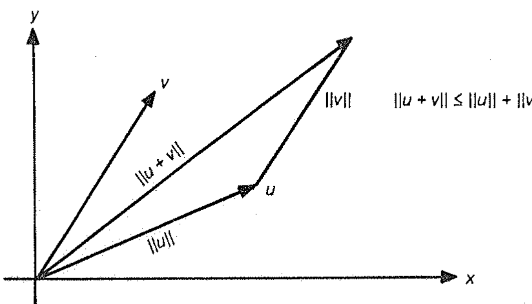
como se quería probar.

Q.E.D.

La desigualdad

$$\|v + u\| \leq \|v\| + \|u\|$$

de la propiedad (3) del teorema anterior, se conoce como *desigualdad triangular*, pues si $V = \mathbf{R}^2$ y se toma en él el producto interno canónico, esta desigualdad establece el conocido hecho geométrico de que en un triángulo cualquiera, la longitud de uno de sus lados no puede exceder la suma de las longitudes de los otros dos lados.



Se introduce ahora el concepto de "distancia entre dos vectores" en un espacio con producto interno.

Sean v y u dos vectores en el espacio V con producto interno $(\cdot \mid \cdot)$, defínase *la distancia* de v a u, denotada por $d(v, u)$, como

$$d(v, u) = \|v - u\|$$

Obsérvese nuevamente que esta definición hace perfecto sentido en cualquier espacio vectorial con producto interno. Si se consideran los casos particulares $V = \mathbf{R}^2$ o $\mathbf{R}^3$ con el producto interno canónico, se ve que $d(v, u)$ no es más que la distancia natural entre los puntos (los vectores) v y u.

Por ejemplo, si $v = (x_1, y_1, z_1)$ y $u = (x_2, y_2, z_2)$ son dos vectores en $\mathbf{R}^3$ y se considera el producto interno canónico de este espacio, se tiene

$$d(v, u) = \|v - u\| = \|(x_1 - x_2, y_1 - y_2, z_1 - z_2)\|$$
$$= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

El concepto de distancia entre dos vectores en un espacio con producto interno generaliza entonces la idea natural de distancia entre dos puntos en el espacio $\mathbf{R}^2$ y $\mathbf{R}^3$.

EJEMPLO 2 Por ejemplo, la distancia de la función sen x a la función cos x en el espacio $C([0, 1])$ es:

$$d(\text{sen } x, \cos x) = \|\text{sen } x - \cos x\|$$

$$= \sqrt{\int_0^1 (\text{sen } x - \cos x)^2 dx} = \sqrt{1 - \text{sen}^2 1} \approx 0.54$$

mientras que la distancia, de la función sen x a la función x es

$$d(\text{sen } x, x) = \|\text{sen } x - x\|$$

$$= \sqrt{\int_0^1 \text{sen } x - \cos x)^2\, dx} = \sqrt{\frac{5}{6} - \frac{1}{4}\text{sen } 2 + 2\cos 1 - 2\,\text{sen } 1} \approx 0.0606$$

como $d(\text{sen } x, x) < d(\text{sen } x, \cos x)$, se puede decir que la función cos x se encuentra "más lejos" de sen x que la función x.

En el siguiente teorema se establecen las propiedades más importantes de la distancia entre dos vectores en un espacio con producto interno.

TEOREMA 2.2

> Sea V un espacio vectorial con producto interno $(\cdot \mid \cdot)$, la distancia $d(u, v) = \|u - v\|$ del vector $u \in V$ al vector $v \in V$, goza de las siguientes propiedades:
>
> 1) $d(u, v) \geq 0 \quad \forall\, u, v \in V, \qquad d(u, v) = 0 \Leftrightarrow u = v$
> 2) $d(u, v) = d(v, u) \quad \forall\, u, v \in V$
> 3) $d(u, v) \leq d(u, w) + d(w, v) \qquad \forall\, u, v, w \in V.$

DEMOSTRACIÓN El hecho de que $d(u, v) \geq 0\ \forall\, u, v \in V$, es una consecuencia inmediata de la propiedad 1) del teorema 2.1.

También por esta misma propiedad se deduce que $d(u, v) = \|u - v\| = 0$ si, y sólo si $u - v = 0$, esto es, $u = v$.

Ahora,

$$d(u, v) = \|u - v\| = \|-(v - u)\| = |-1|\ \|v - u\| = \|v - u\| = d(v, u)$$

Por último, al usar la propiedad 3) del teorema 2.1 se prueba

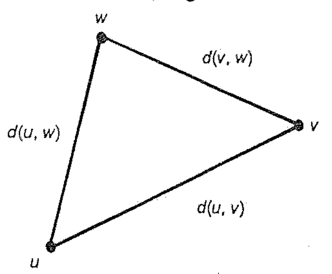
$$d(u, v) = \|u - v\| = \|(u - w) + (w - v)\| \leq \|u - w\| + \|w - v\| = d(u, w) + d(w, v)$$

Q.E.D.

Las propiedades 1), 2) y 3) del teorema anterior son en realidad las propiedades que se exigirían a cualquier concepto sensato de "distancia": la propiedad 1) dice que siempre que se mida una distancia entre dos vectores, dará por resultado un número no negativo y que la única manera de obtener el valor cero en ella, es cuando se mida la distancia de un vector a él mismo.

La propiedad 2) dice que da lo mismo medir la distancia de v a u que de u a v.

Por último, la propiedad 3) es una forma un poco más general de la desigualdad triangular presentada en el teorema 2.1, según se muestra en la siguiente figura:



La desigualdad triangular es también una propiedad que naturalmente se esperaría del concepto de distancia.

EJERCICIOS (SECCIÓN 2, CAPÍTULO 5)

1. Considere el espacio $\mathbf{R}^3$ con el producto interno canónico.
   Calcule las normas de los siguientes vectores:
   a) $(1, 0, 0)$
   b) $(0, 1, 0)$
   c) $\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$
   d) $(3, 1, 4)$
   e) $(-1, 0, 0)$
   f) $(0, -3, -1)$
2. En $\mathbf{R}^3$ considere el producto interno
   $$(x \mid y) = x_1y_1 + 2x_2y_2 + 3x_3y_3$$
   Calcule las normas de los vectores de $\mathbf{R}^3$ del ejercicio anterior.
3. En $M_{2\times 2}$ considere el producto interno canónico. Calcule la norma de los siguientes vectores:
   a) $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$
   b) $\begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$
   c) $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
   d) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
   e) $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$
   f) $\begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$

4. En $C([-1, 1])$ considere el producto interno

$$(f \mid g) = \int_{-1}^{1} f(x)g(x)dx$$

Calcule las normas de los siguientes vectores:

a) e^x
b) $\operatorname{sen} x$
c) x
d) x^2
e) $\cos x$
f) x^2e^x

5. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$, se dice que el vector $v \in V$ es un vector *unitario* si $\|v\| = 1$. Demuestre que los vectores de la base canónica de

a) $\mathbf{R}^n$,
b) P_n,
c) $M_{m\times n}$

con el producto interno canónico son vectores unitarios.

6. Sea V un espacio con el producto interno $(\cdot \mid \cdot)$. Compruebe que la norma de un vector $v \in V$ es la misma que la norma del inverso aditivo del vector v.

7. Suponga que el vector v del espacio V con producto interno $(\cdot \mid \cdot)$ tiene por norma 3. Determine la norma en los siguientes vectores en V:

a) $2v$
b) $-2v$
c) $-3v$
d) $-8v$

8. Verifique la desigualdad triangular para los siguientes pares de vectores:

a) $v = (2, 1, 4)$, $u = (3, 1, 2)$ en $\mathbf{R}^3$ con el producto interno canónico.
b) $v = (2, 1, 4)$, $u = (3, 1, 2)$ en $\mathbf{R}^3$ con el producto interno

$$(x \mid y) = x_1y_1 + 2x_2y_2 + 3x_3y_3$$

c) $v = 1 + x - 3x^2$, $u = 2 + 8x + x^2$ en P_2 con el producto interno canónico.
d) $v = \begin{bmatrix} 2 & 1 \\ -3 & 4 \end{bmatrix}$, $u = \begin{bmatrix} -2 & -1 \\ 3 & -4 \end{bmatrix}$ en $M_{2\times 2}$ con el producto interno canónico.
e) $v = x$, $u = -x + 3$ en $C([-1, 1])$ con el producto interno del ejercicio 4.
f) $v = \operatorname{sen} x$, $u = \cos x$ en $C([-1, 1])$ con el producto interno del ejercicio 4.

9. Considere el espacio $\mathbf{R}^3$ con el producto interno canónico.
Sean $v = (1, 3, -4)$, $u = (3, 9, -12) = 3v$,

a) calcule la norma de los vectores v, u, $v + u$,
b) verifique que en este caso se cumple la igualdad de la desigualdad triangular.

10. Considere el espacio $\mathbf{R}^3$ con el producto interno

$$(x \mid y) = 3x_1y_1 + 8x_2y_2 + 2x_3y_3$$

Considere nuevamente los vectores v y u del ejercicio anterior, ¿se cumple la igualdad de la desigualdad triangular en este caso?

11. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$. Demuestre que se cumple la igualdad de la desigualdad triangular si, y sólo si uno de los vectores es un múltiplo *no negativo* del otro vector. Interprete este hecho geométricamente en los casos de $\mathbf{R}^2$ y $\mathbf{R}^3$ con el producto interno canónico.

12. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$. Pruebe que para cualesquiera dos vectores v y u en V se cumple que

$$\big|\, \|v\| - \|u\| \,\big| \leq \|v - u\|$$

13. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$. Sean v y u dos vectores cualesquiera de V, demuestre que

$$\|v + u\|^2 + \|v - u\|^2 = 2(\|v\|^2 + \|u\|^2)$$

Esta expresión es conocida como "ley del paralelogramo". ¿Cuál es el contenido geométrico de esta expresión en el caso del espacio vectorial $\mathbf{R}^2$ con el producto interno canónico?

14. Calcule la distancia entre los siguientes pares de vectores de $\mathbf{R}^2$ con el producto interno canónico.

a) $v = (1, 3)$, $u = (2, 1)$
b) $v = (-1, 0)$, $u = (0, -1)$
c) $v = (2, 1)$, $u = (-2, -1)$

15. En el espacio $M_{2\times 2}$ con el producto interno canónico, sea

$$v = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

Calcule la distancia del vector v al vector u si

a) $u = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
b) $u = \begin{bmatrix} 3 & -1 \\ 2 & -1 \end{bmatrix}$
c) $u = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$

16. En el espacio $C([-1, 1])$ considere el producto interno

$$(f \mid g) = \int_{-1}^{1} f(x)g(x)dx$$

Sea $f(x) = e^x$, $g(x) = x$, $h(x) = -x^2 + 1$. De los vectores $g(x)$ y $h(x)$ en $C([-1, 1])$. ¿Cuál se encuentra más cerca de $f(x)$?

17. Considere el espacio $C([-1, 1])$ con el producto interno del ejercicio anterior.
Sea $f(x) = x^2 + 1$, $g(x) = x + \frac{3}{4}$, $h(x) = -e^x - 2$, haga una gráfica en la que se muestren (en el intervalo $[-1, 1]$) las funciones $f(x)$, $g(x)$ y $h(x)$. ¿Podría decidir cuál de las funciones $g(x)$ o $h(x)$ se encuentra más cerca de $f(x)$ viendo solamente la gráfica de estas funciones? Verifique su conjetura calculando la distancia de $g(x)$ y $h(x)$ a $f(x)$.

② 18. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)_1$. Sean v, u, w tres vectores en V tales que

$$d(v, u) = d(v, w)$$

Suponga ahora que en el mismo espacio V se considera *otro* producto interno, dígase $(\cdot \mid \cdot)_2$. ¿Puede acontecer que con este nuevo producto interno se tenga que

$$d(v, u) \neq d(v, w)?$$

② 19. Sea V un espacio vectorial con el producto interno $(\cdot|\cdot)_1$. Sean v, u, w tres vectores en V tales que

$$d(v, u) > d(v, w)$$

Suponga ahora que en el mismo espacio V se considera *otro* producto interno, dígase $(\cdot|\cdot)_2$. ¿Puede acontecer que con este nuevo producto interno se tenga que

$$d(v, u) \leq d(v, w)?$$

③ 20. En este ejercicio se considera un tipo más general de espacios vectoriales que los espacios vectoriales con producto interno llamados *espacios vectoriales normados.*

Definición: Sea V un espacio vectorial. Una *norma* en V es una función $\|\cdot\|: V \to \mathbf{R}$ que asocia a cada vector $v \in V$ un número real $\|v\|$, llamado la *norma de v*, de modo que se satisfacen las siguientes tres propiedades:

1) $\|v\| \geq 0 \quad \forall\, v \in V$ y $\|v\| = 0$ si, y sólo si $v = 0$
2) $\|cv\| = |c|\,\|v\| \quad c \in \mathbf{R}, \quad \forall\, v \in V$
3) $\|v + u\| \leq \|v\| + \|u\| \quad \forall\, v, u \in V$

Un espacio vectorial V en el que está definida una norma se llama *espacio vectorial normado.*

a) Demuestre que si V es un espacio vectorial con el producto interno $(\cdot|\cdot)$ entonces V es automáticamente un espacio vectorial normado con la norma $\|\cdot\| = \sqrt{(\cdot|\cdot)}$.

b) En el espacio vectorial $\mathbf{R}^n$, defina la función $\|\cdot\|_1: \mathbf{R}^n \to \mathbf{R}$ como

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

en donde $x = (x_1, x_2, \ldots, x_n)$. Demuestre que esta función es una norma en $\mathbf{R}^n$ (llamada *norma de la suma*).

c) En el espacio vectorial $\mathbf{R}^n$, defina la función $\|\cdot\|_2: \mathbf{R}^n \to \mathbf{R}$ como

$$\|x\|_2 = \text{máx}\,\{|x_i|,\ i = 1, 2, \ldots, n\}$$

en donde $x = (x_1, x_2, \ldots, x_n)$. Compruebe que esta función es una norma en $\mathbf{R}^n$ (llamada *norma del máximo*).

d) En referencia a los dos incisos anteriores, calcule las normas $\|\cdot\|_1$ y $\|\cdot\|_2$ de los siguientes vectores:

d.1) (1, 0, 0)	d.5) (0, 0, 3, 0, 0, 0)
d.2) (2, 3, 5, −1)	d.6) (1, 6, −1, −1, 0, 2)
d.3) (−1, −1, −1, −1)	d.7) (3, 2, 8, −1, 0, 0, 0)
d.4) (−20, 5, 2, 4)	d.8) (0, 0, 2, 0, 0, 3, 0)

e) En el espacio vectorial $C([-1, 1])$ defina la función $\|\cdot\|: C([-1, 1]) \to \mathbf{R}$ como

$$\|f\| = \underset{x \in [-1, 1]}{\text{máx}}\, |f(x)|$$

Demuestre primeramente que ésta es una buena definición para esta función (recuerde que toda función continua que está definida en un intervalo cerrado finito alcanza su máximo dentro del intervalo).

Pruebe que esta función define una norma en $C([-1, 1])$.

Calcule las normas de los siguientes vectores en $C([-1, 1])$:

e.1) $f(x) = x^2 + 1$	e.5) $f(x) = e^x$
e.2) $f(x) = -x^2 + 1$	e.6) $f(x) = x^3 - x^2 + x + 1$
e.3) $f(x) = x$	e.7) $f(x) = x^3 + 3x^2 + 4$
e.4) $f(x) = \text{sen}\, x$	e.8) $f(x) = x^5 + x^3 + 3x$

f) Sea V un espacio vectorial con el producto interno $(\cdot|\cdot)$. En el inciso a) se demostró que V es un espacio vectorial normado con la norma $\|\cdot\| = \sqrt{(\cdot|\cdot)}$. En este caso, se dice que *la norma* $\|\cdot\|$ *proviene del producto interno* $(\cdot|\cdot)$. Demuestre que si la norma $\|\cdot\|$ proviene del producto interno $(\cdot|\cdot)$, entonces se cumple la *ley del paralelogramo*

$$\|v + u\|^2 + \|v - u\|^2 = 2(\|v\|^2 + \|u\|^2) \quad \forall\, v, u \in V$$

g) Compruebe que la norma $\|\cdot\|_1$ en $\mathbf{R}^n$ del inciso b) no proviene de un producto interno en $\mathbf{R}^n$.

(*Sugerencia*: Considere en $\mathbf{R}^2$ los vectores (1, 0) y (0, 1). Use el inciso anterior.)

h) Demuestre que la norma $\|\cdot\|_2$ en $\mathbf{R}^n$ del inciso c) no proviene de un producto interno en $\mathbf{R}^n$.

i) Compruebe que la norma $\|\cdot\|$ en $C([-1, 1])$ del inciso e) no proviene de un producto interno en $C([-1, 1])$.

j) En el inciso f) se ha demostrado que una condición *necesaria* para que la norma $\|\cdot\|$ en el espacio vectorial V provenga de un producto interno en V es que se cumpla la ley del paralelogramo. Se puede demostrar que esta condición también es *suficiente* (la demostración es un tanto complicada).

En $\mathbf{R}^3$ considere la función $\|\cdot\|: \mathbf{R}^3 \to \mathbf{R}$ dada por

$$\|(x, y, z)\| = \sqrt{3x^2 + y^2 + 4z^2}$$

Demuestre que esta función es una norma en $\mathbf{R}^3$ y que ella proviene de un producto interno en $\mathbf{R}^3$.

③ 21. En este ejercicio se considera un tipo más general de espacios que los espacios vectoriales normados introducidos en el ejercicio anterior, llamados "*espacios métricos*".

Definición. Sea M un conjunto no vacío. Una *métrica* en M es una función $d: M \times M \to \mathbf{R}$ que asocia a cada par de elementos x, y en M un número real $d(x, y)$ llamado *distancia de x a y*, de modo que se satisfacen las siguientes tres propiedades:

1) $d(x, y) \geq 0\ \forall\, x, y \in M \quad d(x, y) = 0 \Leftrightarrow x = y.$
2) $d(x, y) = d(y, x) \quad \forall\, x, y \in M$
3) $d(x, z) \leq d(x, y) + d(y, z) \quad \forall\, x, y, z \in M$

A un conjunto M en el que está definida una métrica $d: M \times M \to \mathbf{R}$ se le llama *espacio métrico.*

a) Demuestre que si M es un espacio vectorial normado, entonces M es automáticamente un espacio métrico con la métrica $d(x, y) = \|x - y\|$, en donde $\|\cdot\|$ es la norma de M. Concluya entonces que todo espacio vectorial con producto interno es un espacio métrico.

b) Defina la función $d: M \times M \to \mathbf{R}$ como

$$d(x, y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{si } x \neq y \end{cases}$$

Demuestre que esta función es una métrica en M.

c) Defina la función $d_1: \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ como

$$d_1(x, y) = \sum_{i=1}^{n} |x_i - y_i|$$

en donde $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$. Compruebe que d_1 es una métrica en $\mathbf{R}^n$.

(*Sugerencia*: véase el inciso b) del ejercicio anterior.)

d) Defina la función $d_2: \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ como

$$d_2(x, y) = \text{máx}\,\{|x_i - y_i|, \quad i = 1, 2, \ldots, n\}$$

Demuestre que d_2 es una métrica en $\mathbf{R}^n$.

e) Defina la función $d: C([-1, 1]) \times C([-1, 1]) \to \mathbf{R}$ como

$$d(f, g) = \max_{x \in [-1, 1]} |f(x) - g(x)|$$

Compruebe que d es una métrica en $C([-1, 1])$.

f) Defina la función $d: \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ como

$$d(x, y) = (x - y)^2$$

Demuestre que esta función *no* es una métrica en $\mathbf{R}$.

g) A partir de una métrica $d: M \times M \to \mathbf{R}$ en el conjunto M, se pueden generar otras métricas en M definiendo

$$d_1(x, y) = \sqrt{d(x, y)}$$

$$d_2(x, y) = \frac{d(x, y)}{1 + d(x, y)}$$

$$d_3(x, y) = \text{mín}\,\{1, d(x, y)\}$$

Pruebe que d_1, d_2 y d_3 son también métricas en M.

h) Sea V un espacio vectorial normado. En el inciso a) se demostró que V es un espacio métrico definiendo la métrica $d: V \times V \to \mathbf{R}$ como $d(x, y) = \|x - y\|$. En este caso, se dice *que la métrica* d *proviene de la norma* $\|\cdot\|$. Demuestre que una condición necesaria y suficiente para que una métrica d en V provenga de una norma $\|\cdot\|$ en V es que para todo $v, u \in V$ y $c \in \mathbf{R}$ se cumpla

$$d(v + w, u + w) = d(v, u),\ w \in V$$

$$d(cv, cu) = |c|\, d(v, u)$$

i) Si M es un espacio vectorial. Demuestre que la métrica en M del inciso b) no proviene de una norma.

j) Concluya de este ejercicio que:

j.1) Todo espacio con producto interno es un espacio normado.

j.2) Todo espacio normado es un espacio métrico.

j.3) Todo espacio con producto interno es un espacio métrico.

Dé ejemplos que muestren que las afirmaciones recíprocas son falsas.

③ 22. El objetivo de este ejercicio es establecer dos desigualdades muy importantes en matemáticas conocidas como desigualdad de Hölder y desigualdad de Minkowski.

a) Sea $p \in \mathbf{R}$, $p > 1$. Defina el número $q \in \mathbf{R}$ como el número real tal que $p^{-1} + q^{-1} = 1$. Se dirá que p y q son exponentes conjugados. Demuestre que si p y q son exponentes conjugados entonces $(p - 1)(q - 1) = 1$. Concluya entonces que si $y = x^{p-1}$, entonces $x = y^{q-1}$, en donde p y q son exponentes conjugados.

b) Sean p y q dos exponentes conjugados, considere la función $f: [0, a] \to \mathbf{R}$, $f(x) = x^{p-1}$. Use argumentos geométricos para establecer la desigualdad

$$ab \le \int_0^a x^{p-1}dx + \int_0^b y^{q-1}dy$$

en donde b es un número real no negativo cualquiera. (*Sugerencia*: construya una gráfica en la que se muestre la función $y = x^{p-1}$. Dibuje el intervalo $[0, b]$ sobre el eje y. Interprete ambos miembros de la desigualdad anterior como áreas).

c) Concluya del resultado del ejercicio anterior que si a y b son dos números reales no negativos cualesquiera y p y q son dos exponentes conjugados, entonces

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}$$

d) Sean $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$ y $\tilde{y} = (\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_n)$ dos vectores de $\mathbf{R}^n$ tales que

$$\sum_{i=1}^{n} |\tilde{x}_i|^p = 1 \qquad \sum_{i=1}^{n} |\tilde{y}_i|^q = 1$$

en donde p y q son dos exponentes conjugados. Use el resultado del inciso anterior para demostrar que para cualquier $i = 1, 2, \ldots, n$ se tiene

$$|\tilde{x}_i \tilde{y}_i| \le \frac{|\tilde{x}_i|^p}{p} + \frac{|\tilde{y}_i|^q}{q}$$

y concluya entonces (sumando desde $i = 1$ hasta n) que

$$\sum_{i=1}^{n} |\tilde{x}_i \tilde{y}_i| = 1$$

e) Demuestre la *desigualdad de Hölder*

$$\sum_{i=1}^{n} |x_i y_i| \le \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}$$

en donde $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ son números reales cualesquiera y p y q son exponentes conjugados.

(*Sugerencia*: considere dos vectores cualesquiera de $\mathbf{R}^n$ $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$. Defina los vectores $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$, $\tilde{y} = (\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_n)$ como

$$\tilde{x}_i = \frac{x_i}{\left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}} \qquad \tilde{y}_i = \frac{y_i}{\left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}} \qquad i = 1, 2, \ldots, n$$

(Aplique ahora el resultado del inciso anterior.)

f) ¿A qué equivale la desigualdad de Hölder cuando $p = q = 2$?

g) Sean $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ dos vectores cualesquiera de $\mathbf{R}^n$. Demuestre que para cualquier $i = 1, 2, \ldots, n$ se tiene

$$|x_i + y_i|^p \leq (|x_i| + |y_i|)\, |x_i + y_i|^{p-1}$$

de donde $p \in \mathbf{R}$, $p > 1$. Concluya entonces, sumando desde $i = 1$ hasta n que

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq \sum_{i=1}^{n} |x_i|\, |x_i + y_i|^{p-1} + \sum_{i=1}^{n} |y_i|\, |x_i + y_i|^{p-1}$$

h) Aplique la desigualdad de Hölder a cada una de las sumatorias que aparecen en el miembro derecho de la desigualdad establecida en el inciso anterior para demostrar que si p y q son exponentes conjugados entonces

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq \left[\left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}\right]\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/q}$$

i) Concluya del resultado del inciso anterior la *desigualdad de Minkowski*

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}$$

en donde $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ son números reales cualesquiera y $p \in \mathbf{R}$, $p > 1$.

j) ¿A qué equivale la desigualdad de Minkowski cuando $p = 2$?

③ 23. Una sucesión de números reales es una función del tipo f: $\mathbf{N} \to \mathbf{R}$. Denotando a $f(i)$ como a_i, $i = 1, 2, \ldots$, se escribirá la sucesión f: $\mathbf{N} \to \mathbf{R}$ como $(a_1, a_2, \ldots, a_n, \ldots)$ o bien, más abreviadamente como (a_i). Sea p un real mayor o igual a 1 (arbitrario pero fijo). Denote por l^p (se lee "ele pe") al conjunto de todas las sucesiones de números reales (a_i) tales que

$$\sum_{i=1}^{\infty} |a_i|^p < \infty$$

Nuestro interés en este ejercicio es demostrar que el conjunto l^p con ciertas operaciones de suma y de producto por escalares que se definirán más adelante es un espacio

vectorial, y estudiar algunas propiedades importantes de este espacio para ciertos valores concretos de p.

a) Primeramente, extienda la desigualdad de Minkowski establecida en el inciso i) del ejercicio anterior para el caso de sumas infinitas. Es decir, demuestre que

$$\left(\sum_{i=1}^{\infty} |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{\infty} |y_i|^p\right)^{1/p}$$

en donde x_i, y_i, $i = 1, 2, \ldots$ son números cualesquiera.

b) En el conjunto l^p defina la suma de la sucesión $(a_i) \in l^p$ y la sucesión $(b_i) \in l^p$ como

$$(a_i) + (b_i) = (a_i + b_i)$$

y si $k \in \mathbf{R}$, defina el producto del escalar k por la sucesión $(a_i) \in l^p$ como

$$k(a_i) = (ka_i)$$

Demuestre que si (a_i) y (b_i) son sucesiones de l^p entonces $(a_i) + (b_i)$ y $k(a_i)$ son también sucesiones de l^p.
(*Sugerencia*: para demostrar que la suma de sucesiones de l^p es una sucesión de l^p, use el resultado del inciso anterior.)

c) Compruebe que el conjunto l^p con las operaciones de suma y producto por escalares definidas en el inciso anterior es un espacio vectorial. ¿Cuál es el vector cero de este espacio?, ¿cuál es el inverso aditivo de una sucesión (a_i) de l^p?

d) En el espacio l^p defina la función $\|\cdot\|_p$: $l^p \to \mathbf{R}$ como

$$\|(a_i)\|_p = \left(\sum_{i=1}^{\infty} |a_i|^p\right)^{1/p}$$

d.1) Demuestre que $\|(a_i)\|_p \geq 0\ \forall\ (a_i) \in l^p$ y que $\|(a_i)\|_p = 0$ si, y sólo si $(a_i) = 0$ (el vector cero de l^p).

d.2) Compruebe que si $k \in \mathbf{R}$ y $(a_i) \in l^p$, entonces $\|k(a_i)\|_p = |k|\, \|(a_i)\|_p$.

d.3) Demuestre que si $(a_i), (b_i) \in l^p$, entonces

$$\|(a_i) + (b_i)\|_p \leq \|(a_i)\|_p + \|(b_i)\|_p$$

(Sugerencia: use el resultado del inciso a).)

d.4) Concluya entonces que la función $\|\cdot\|_p$ es una norma en el espacio l^p y que por tanto, l^p con esta norma es un espacio vectorial normado.

e) Demuestre que el espacio l^p no es de dimensión finita.

f) Suponga que V es un espacio vectorial normado con la norma $\|\cdot\|$ y que existe una sucesión de vectores en V, diga $v_1, v_2, \ldots, v_n, \ldots$ tales que para cualquier vector $v \in V$ existe una sucesión de números reales $c_1, c_2, \ldots, c_n, \ldots$ (bien determinada por el vector v) con la propiedad

$$\lim_{n \to \infty} \|v - (c_1v_1 + c_2v_2 + \ldots + c_nv_n)\| = 0$$

Se dice entonces que el conjunto (infinito) $\{v_1, v_2, \ldots, v_n, \ldots\}$ es una *base* (llamada

base de Schauder) del espacio normado V. En tal caso, se puede escribir

$$v = \sum_{i=1}^{\infty} c_i v_i$$

En el espacio normado l^p con la norma $\|\cdot\|_p$ considere los vectores

$$\begin{aligned} e_1 &= (1, 0, 0, 0, \ldots) \\ e_2 &= (0, 1, 0, 0, \ldots) \\ e_3 &= (0, 0, 1, 0, \ldots) \\ &\vdots \end{aligned}$$

Es decir, e_i es el vector de l^p, (a_j) diga, tal que $a_j = 0$ si $j \neq i$ y $a_i = 1$. Pruebe que estos vectores constituyen una base del espacio l^p. Más aún, demuestre que si $(a_i) \in l^p$, entonces

$$(a_i) = \sum_{i=1}^{\infty} a_i e_i$$

g) Considere el espacio l^2. Demuestre que la norma $\|\cdot\|_2$ definida anteriormente

$$\|(a_i)\|_2 = \left(\sum_{i=1}^{\infty} |a_i|^2 \right)^{1/2}$$

proviene de un producto interno en l^2 (véanse incisos f) y j) del ejercicio 20 de esta sección). Más aún, demuestre que tal producto interno es $(\cdot \mid \cdot)\colon l^2 \times l^2 \to \mathbf{R}$

$$((a_i) \mid (b_i)) = \sum_{i=1}^{\infty} a_i b_i$$

h) Considere ahora el espacio l^p con $p \neq 2$, con la norma $\|\cdot\|_p$ definida en el inciso d). Tome los vectores $v_1 = (1, 1, 0, 0, \ldots)$ y $v_2 = (1, -1, 0, 0, \ldots)$ (observe que v_1 y v_2 son vectores l^p). Demuestre que

$$\begin{aligned} \|v_1 + v_2\| = \|v_1 - v_2\| = 2 \\ \|v_1\| = \|v_2\| = 2^{1/p} \end{aligned}$$

i) Use los resultados del inciso anterior para demostrar que la norma $\|\cdot\|_p$ del espacio l^p, $p \neq 2$, *no* proviene de un producto interno en l^p (véase inciso f) del ejercicio 20 de esta sección).

② 24) En el espacio vectorial l^p considere los operadores $T_1, T_2\colon l^p \to l^p$ definidos como

$$\begin{aligned} T_1[(a_1, a_2, a_3, \ldots)] &= (a_2, a_3, \ldots) \\ T_2[(a_1, a_2, a_3, \ldots)] &= (0, a_1, a_2, a_3, \ldots) \end{aligned}$$

a) Demuestre que T_1 y T_2 son operadores lineales en l^p.

b) Compruebe que T_1 es sobreyectivo pero no inyectivo.

c) Demuestre que T_2 es inyectivo pero no sobreyectivo.

d) Pruebe que $T_1 \circ T_2$ es el operador identidad en l^p (y por lo tanto es inyectivo y sobreyectivo).

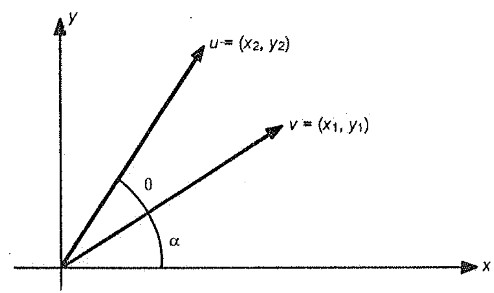
e) Demuestre que $T_2 \circ T_1$ no es el operador identidad en l^p.

f) ¿Es T_1 un operador inversible?, ¿lo es T_2?

[**Nota:** los incisos b), c) y d) dan una respuesta a los ejercicios 2 y 3 de la sección 6 del capítulo 4.]

3. ÁNGULO ENTRE DOS VECTORES. ORTOGONALIDAD

En el espacio vectorial $\mathbf{R}^2$ se tiene bien definida la noción de ángulo θ entre dos vectores $v, u, \in \mathbf{R}^2$



En efecto, usando un argumento elemental de geometría analítica se puede establecer el valor de θ:

pendiente de la recta en la que se encuentra $u = \tan(\theta + \alpha) = \dfrac{y_2}{x_2}$

pendiente de la recta en la que se encuentra $v = \tan \alpha = \dfrac{y_1}{x_1}$

Entonces

$$\begin{aligned} \theta &= \arctan \frac{y_2}{x_2} - \arctan \frac{y_1}{x_1} \\ &= \arctan \left(\frac{x_1 y_2 - x_2 y_1}{x_2 x_1 + y_2 y_1} \right) \end{aligned}$$

Al expresar la fórmula anterior en términos de la función coseno se obtiene

$$\cos \theta = \frac{x_1 x_2 + y_1 y_2}{(x_1^2 + y_1^2)(x_2^2 + y_2^2)} \qquad (3.1)$$

Obsérvese que tanto en el numerador como en el denominador de la fórmula (3.1) aparecen expresiones familiares, a saber

$$\begin{aligned} (v \mid u) &= x_1 x_2 + y_1 y_2 \\ \|v\|^2 &= x_1^2 + y_1^2 \\ \|u\|^2 &= x_2^2 + y_2^2 \end{aligned}$$

en donde $(\cdot\mid\cdot)$ es el producto interno canónico de $\mathbb{R}^2$, de modo que (3.1) puede ser reescrito como

$$\cos\theta = \frac{(v\mid u)}{\|v\|\,\|u\|} \tag{3.2}$$

Nada impide interpretar la fórmula (3.2) independientemente del análisis previo a ella, y usarla para *definir* el ángulo entre dos vectores v y u en cualquier espacio vectorial V con producto interno $(\cdot\mid\cdot)$. Esto es precisamente lo que se hará ahora.

3.1. ÁNGULO ENTRE DOS VECTORES

Sea $(V, (\cdot\mid\cdot))$ un espacio con producto interno y sean v, u dos vectores no nulos de V.

La desigualdad de Cauchy-Schwartz establece que

$$|(v\mid u)| \leq \|v\|\,\|u\|$$

de donde

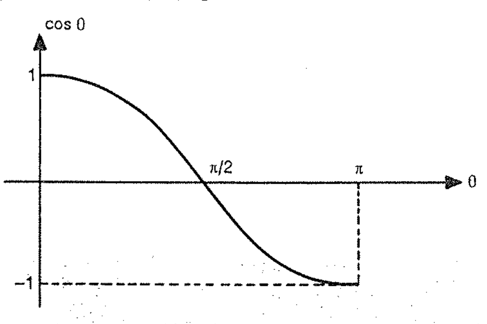
$$-1 \leq \frac{(v\mid u)}{\|v\|\,\|u\|} \leq 1$$

La función

$$f(v, u) = \frac{(v\mid u)}{\|v\|\,\|u\|}$$

es entonces una función cuyo rango es $[-1, 1]$.

Obsérvese que la función $\cos\theta$, es una función biyectiva en el intervalo $0 \leq \theta \leq \pi$ cuyo rango es precisamente $[-1, 1]$.



Inspirados por la fórmula (3.2) establézcase entonces la siguiente definición:

DEFINICIÓN 3.1 Si v y u son dos vectores no nulos en el espacio vectorial V con producto interno $(\cdot\mid\cdot)$, defínase el *ángulo* $(0 \leq \theta \leq \pi)$ entre v y u como

$$\theta = \arccos\frac{(v\mid u)}{\|v\|\,\|u\|}$$

EJEMPLO 1 Así por ejemplo, si $V = P_3$ con el producto interno canónico, el ángulo entre el polinomio

$$p = 1 + 3x - 2x^2 + x^3$$

y el polinomio

$$q = -2 + x - x^2 + x^3$$

es

$$\theta = \arccos\left(\frac{(p\mid q)}{\|p\|\,\|q\|}\right)$$

$$= \arccos\left(\frac{(1)(-2) + (3)(1) + (-2)(-1) + (1)(1)}{((1)^2 + (3)^2 + (-2)^2 + (1)^2)((-2)^2 + (1)^2 + (-1)^2 + (1)^2)}\right)$$

$$= \arccos\left(\frac{4}{105}\right) \approx 87.817°$$

En un espacio con producto interno se tiene la versión abstracta de la ley de los cosenos.

TEOREMA 3.1 (LEY DE LOS COSENOS.) Si v y u son dos vectores en un espacio con producto interno, y θ es el ángulo entre v y u, entonces

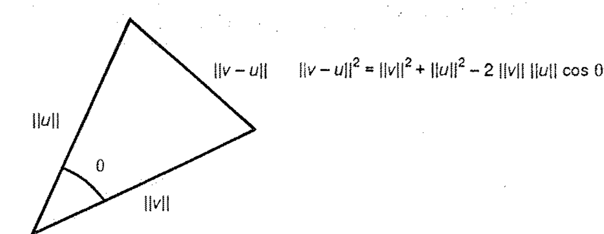
$$\|v - u\|^2 = \|v\|^2 + \|u\|^2 - 2\|v\|\,\|u\|\cos\theta$$

DEMOSTRACIÓN Es una consecuencia directa de la definición de norma de un vector y de la fórmula (3.2).

$$\|v - u\|^2 = (v - u\mid v - u) = (v\mid v) - 2(v\mid u) + (u\mid u)$$
$$= \|v\|^2 + \|u\|^2 - 2\|v\|\,\|u\|\cos\theta$$

Q.E.D.

Claramente, si V es el espacio $\mathbb{R}^2$ con el producto interno canónico, el teorema 3.1 establece la "versión concreta" (la que se conoce) de la ley de los cosenos



En el caso concreto del espacio $\mathbf{R}^2$ con el producto interno canónico, se sabe que una condición equivalente al hecho de que los vectores v y u sean perpendiculares es que su producto interno (que en este caso coincide con su producto punto) sea cero.

Esta equivalencia se puede también deducir de la fórmula (3.2): si v y u son dos vectores perpendiculares no nulos en $\mathbf{R}^2$, el ángulo que forman entre ellos es $\theta = \frac{\pi}{2}$, de modo que

$$0 = \cos\frac{\pi}{2} = \frac{(v \mid u)}{\|v\|\,\|u\|}$$

de donde

$$(v \mid u) = 0$$

Similarmente, si $(v \mid u) = 0$ entonces

$$\cos\theta = \frac{(v \mid u)}{\|v\|\,\|u\|} = 0 \quad \Rightarrow \theta = \pi/2$$

La condición geométrica de perpendicularidad entre los vectores v y u del espacio $\mathbf{R}^2$ se puede generalizar entonces a un espacio vectorial abstracto gracias a la definición 3.1: dos vectores v y u en un espacio vectorial V con producto interno son "perpendiculares" si el ángulo que forman entre ellos es 90°, o equivalentemente.

DEFINICIÓN 3.2 Sea $(V, (\cdot \mid \cdot))$ un espacio con producto interno. Los vectores v y u en V son *ortogonales* si

$$(v \mid u) = 0$$

NOTA: En el caso general, se referirá siempre a la propiedad geométrica de *perpendicularidad* entre dos vectores, como (la propiedad de) *ortogonalidad* entre tales vectores, reservando entonces "perpendicularidad" para el caso concreto del espacio $\mathbf{R}^2$ (o $\mathbf{R}^3$) con el producto interno canónico.

Obsérvese que, según la definición 3.2, el vector cero de V es ortogonal a todo vector $v \in V$, pues por el teorema 1.1

$$(0 \mid v) = 0 \qquad \forall\, v \in V.$$

Si V es el espacio $M_{3\times 2}$ con el producto interno canónico, los vectores

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 3 \\ -2 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 5 \\ 0 & -2 \\ 1 & 2 \end{bmatrix}$$

son ortogonales, pues

$$(A \mid B) = (2)(3) + (0)(5) + (1)(0) + (3)(-2) + (-2)(1) + (1)(2) = 0$$

Similarmente, si V es el espacio $C([0, 1])$, los vectores $f_1 = x$ y $f_2 = 3x - 2$ son ortogonales pues

$$(f_1 \mid f_2) = \int_0^1 f_1(x) f_2(x)dx = \int_0^1 x(3x-2)dx = 0$$

En la siguiente subsección se estudiarán conjuntos de vectores ortogonales en un espacio con producto interno.

3.2. CONJUNTOS ORTONORMALES

Sea V un espacio vectorial con producto interno $(\cdot \mid \cdot)$ y sea S un subconjunto de V. Se dice que S es un *conjunto ortogonal* de vectores en V si cada par de vectores distintos en S son ortogonales.

En particular, si $S = \{v_1, v_2, \ldots, v_k\}$, S será ortogonal si, y sólo si

$$(v_i \mid v_j) = \begin{cases} \|v_i\|^2 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Si además cada vector v del conjunto ortogonal S de V tiene norma 1, al conjunto S se le llama *conjunto ortonormal*.

NOTA: A un vector v del espacio con producto interno V que tiene norma 1 se le llama *vector unitario*. Un conjunto ortonormal de vectores en V es entonces un conjunto ortogonal de vectores unitarios en V.

EJEMPLO 2 Por ejemplo, en el espacio $\mathbf{R}^3$ con el producto interno canónico los vectores

$$e_1 = (1, 0, 0)$$
$$e_2 = (0, 1, 0)$$
$$e_3 = (0, 0, 1)$$

forman un conjunto ortonormal, pues

$$(e_1 \mid e_2) = (e_1 \mid e_3) = (e_2 \mid e_3) = 0$$

(esto es, tales vectores forman un conjunto ortogonal) y además

$$\|e_1\| = \|e_2\| = \|e_3\| = 1$$

EJEMPLO 3 En el espacio vectorial $C([-1, 1])$ de las funciones continuas en el intervalo $[-1, 1]$

con el producto interno $(f \mid g) = \int_{-1}^{1} f(x)g(x)dx$ el conjunto (infinito) de vectores

$$S = \{\cos 2n\pi x,\ n = 0, 1, 2, \ldots\}$$

es un conjunto ortonormal.

En efecto, cada vector $v_n = \cos 2n\pi x$ en S es unitario pues

$$\|v_n\| = \sqrt{(v_n \mid v_n)} = \sqrt{\int_{-1}^{1} \cos^2 2n\pi x\, dx} = 1$$

Por otra parte, si $n \neq m$

$$(v_n \mid v_m) = \int_{-1}^{1} \cos(2n\pi x)\cos(2m\pi x)\, dx$$

$$= \frac{1}{2}\int_{-1}^{1} (\cos 2(n+m)\pi x + \cos 2(n-m)\pi x)\, dx$$

$$= -\frac{1}{2}\left[\frac{1}{2(n+m)\pi}\operatorname{sen} 2(n+m)\pi x + \frac{1}{2(n-m)\pi}\operatorname{sen} 2(n-m)\pi x\right]_{-1}^{1}$$

$$= 0$$

Es decir, los vectores v_n y v_m $(n \neq m)$ son ortogonales.

El siguiente resultado es válido para cualquier conjunto ortogonal (no necesariamente ortonormal) de vectores no nulos en un espacio con producto interno.

TEOREMA 3.2

> Sea S un conjunto ortogonal de vectores no nulos en el espacio V con producto interno $(\cdot \mid \cdot)$. Entonces S es un conjunto linealmente independiente.

DEMOSTRACIÓN Sean $v_1, v_2, \ldots, v_k$ vectores distintos de S. Se debe de probar que

$$c_1v_1 + c_2v_2 + \ldots + c_kv_k = 0 \quad \Leftrightarrow \quad c_1 = c_2 = \ldots = c_k = 0$$

Al tomar el producto interno del vector $0 = c_1v_1 + c_2v_2 + \ldots + c_kv_k$ con el vector v_j, $i \leq j \leq k$ se obtiene

$$0 = (0 \mid v_j) = (c_1v_1 + c_2v_2 + \ldots + c_kv_k \mid v_j) = \left(\sum_{i=1}^{k} c_iv_i \mid v_j\right) = \sum_{i=1}^{k} c_i(v_i \mid v_j)$$

Como el conjunto S es ortogonal, se tiene que $(v_i \mid v_j) \neq 0$ si, y sólo si $i = j$, en cuyo caso se obtiene $(v_i \mid v_j) = \|v_j\|^2$. Entonces,

$$c_j\|v_j\|^2 = 0$$

Como por hipótesis $v_j \neq 0$ (y por tanto, $\|v_j\| \neq 0$), se concluye que

$$c_j = 0$$

Este argumento es válido para toda $j = 1, 2, \ldots, k$. Por tanto, $c_1 = c_2 = \ldots = c_k = 0$, como se quería probar.

Q.E.D.

COROLARIO

> Todo conjunto ortonormal de vectores en un espacio con producto interno es linealmente independiente.

DEMOSTRACIÓN Si S es un conjunto ortonormal, entonces S es ortogonal y ninguno de los vectores de S puede ser cero (pues todos ellos tienen norma 1). El corolario se deduce entonces del teorema anterior.

Q.E.D.

TEOREMA 3.3

> Sea $(V, (\cdot \mid \cdot))$ un espacio con producto interno, y sea W el espacio generado por el conjunto ortonormal de vectores $S = \{v_1, v_2, \ldots, v_k\}$. El vector $u \in V$ pertenece a W si, y sólo si u puede escribirse como
>
> $$u = (u \mid v_1)v_1 + (u \mid v_2)v_2 + \ldots + (u \mid v_k)v_k$$

DEMOSTRACIÓN Si el vector $u \in V$ se escribe como

$$u = (u \mid v_1)v_1 + (u \mid v_2)v_2 + \ldots + (u \mid v_k)v_k$$

es obvio que $u \in W$, pues en tal caso, u es una combinación lineal de los vectores $v_1, v_2, \ldots, v_k$ que generan a W.

Recíprocamente, supóngase que $u \in W$. Entonces existen escalares $c_1, c_2, \ldots, c_k$ tales que

$$u = c_1v_1 + c_2v_2 + \ldots + c_kv_k = \sum_{i=1}^{k} c_iv_i$$

Al tomar el producto interno del vector u con el vector $v_j \in S$, $1 \leq j \leq k$, se tiene

$$(u \mid v_j) = \left(\sum_{i=1}^{k} c_iv_i \mid v_j\right) = \sum_{i=1}^{k} c_i(v_i \mid v_j)$$

Pero $(v_i|v_j) = 0$ si $i \neq j$ y $(v_i \mid v_i) = \| v_i \|^2 = 1$ (pues S es un conjunto ortonormal).

Entonces

$$(u \mid v_j) = c_j$$

por lo que

$$u = \sum_{i=1}^{k} c_i v_i = (u \mid v_1)v_1 + (u \mid v_2)v_2 + \ldots + (u \mid v_k)v_k$$

como se quería demostrar. **Q.E.D.**

El teorema anterior es un resultado muy importante que se usará en la próxima sección. En este momento, se usará para probar la versión generalizada del teorema de Pitágoras.

TEOREMA 3.4

(Teorema de Pitágoras.) Sea $(V, (\cdot \mid \cdot))$ un espacio con producto interno y sea W el subespacio de V generado por el conjunto ortonormal $S = \{v_1, v_2, \ldots, v_k\}$. Si u es un vector de W se tiene

$$\|u\|^2 = \sum_{i=1}^{k} (u \mid v_i)^2$$

DEMOSTRACIÓN Por el teorema anterior se puede escribir

$$u = \sum_{i=1}^{k} (u \mid v_i)v_i$$

Entonces

$$\|u\|^2 = (u \mid u) = \left(\sum_{i=1}^{k} (u \mid v_i)v_i \,\Big|\, u \right)$$

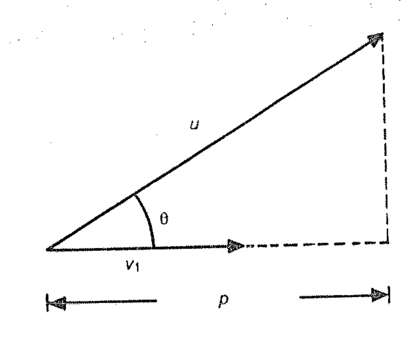
$$= \sum_{i=1}^{k} (u \mid v_i)(v_i \mid u) = \sum_{i=1}^{k} (u \mid v_i)^2$$

Q.E.D.

Interprétese geométricamente el teorema anterior cuando $k = 2$ en el caso en el que V es el espacio $\mathbf{R}^2$ con el producto interno canónico. El teorema establece en tal caso que si $\{v_1, v_2\}$ es un conjunto ortonormal de vectores en $\mathbf{R}^2$, y u es un vector en $\mathcal{L}(v_1, v_2)$ ($=\mathbf{R}^2$, ¿por qué?) entonces

$$\|u\|^2 = (u \mid v_1)^2 + (u \mid v_2)^2$$

De la figura



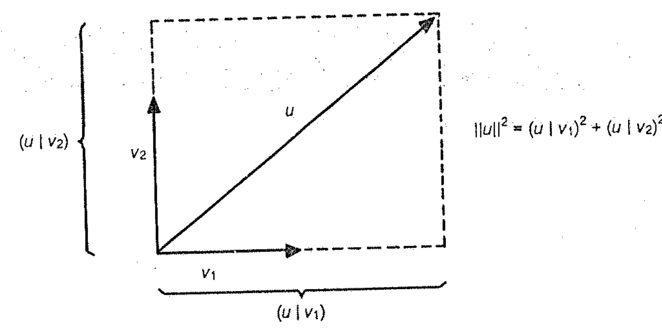
se ve que

$$\cos\theta = \frac{p}{\|u\|}$$

o bien,

$$p = \|u\| \cos\theta$$
$$= \|u\| \, \|v_1\| \cos\theta \quad (\text{pues } \|v_1\| = 1)$$
$$= (u \mid v_1)$$

Entonces se tiene una situación como la mostrada en la siguiente figura:



que no es más que el conocido teorema de Pitágoras: en un triángulo rectángulo la suma de los cuadrados de los catetos ($(u \mid v_1)^2 + (u \mid v_2)^2$) es igual al cuadrado de la hipotenusa ($\|u\|^2$).

Los siguientes dos resultados (la identidad de Parseval y la desigualdad de Bessel) son resultados clásicos sobre conjuntos ortonormales de vectores en espacios con producto interno. Se presentará sólo el caso en el que el conjunto ortonormal involucrado es finito. La importancia de tales resultados es de carácter teórico.

TEOREMA 3.5

(Identidad de Parseval.) Sea $(V, (\cdot \mid \cdot))$ un espacio con producto interno y M el espacio generado por el conjunto ortonormal $S = \{v_1, v_2, \ldots, v_k\}$. Si $u \in M$ y w es cualquier vector de V, se tiene

$$(u \mid w) = \sum_{i=1}^{k} (u \mid v_i)(v_i \mid w)$$

DEMOSTRACIÓN Según el teorema 3.3 se puede escribir el vector $u \in M$ como

$$u = \sum_{i=1}^{k} (u \mid v_i)v_i$$

Entonces

$$(u \mid w) = \left(\sum_{i=1}^{k} (u \mid v_i)v_i \,\middle|\, w \right) = \sum_{i=1}^{k} (u \mid v_i)(v_i \mid w)$$

Q.E.D.

TEOREMA 3.6

(Desigualdad de Bessel.) Sea $(V, (\cdot \mid \cdot))$ un espacio con producto interno, y sea $S = \{v_1, v_2, \ldots, v_k\}$ un conjunto ortonormal de vectores en V. Si u es un vector en V se tiene la desigualdad

$$\|u\|^2 \geq \sum_{j=1}^{k} (u \mid v_j)^2$$

La igualdad en la desigualdad anterior se cumple si, y sólo si u pertenece al espacio generado por S.

DEMOSTRACIÓN Al considerar el vector

$$w = u - \sum_{i=1}^{k} (u \mid v_i)v_i$$

Se tiene

$$\|w\|^2 = (w \mid w) = \left(u - \sum_{i=1}^{k} (u \mid v_i)v_i \,\middle|\, u - \sum_{j=1}^{k} (u \mid v_j)v_j \right)$$

$$= (u \mid u) - \left(u \,\middle|\, \sum_{j=1}^{k} (u \mid v_j)v_j \right) - \left(\sum_{i=1}^{k} (u \mid v_i)v_i \,\middle|\, u \right) + \left(\sum_{i=1}^{k} (u \mid v_i)v_i \,\middle|\, \sum_{j=1}^{k} (u \mid v_j)v_j \right)$$

$$= (u \mid u) - \sum_{j=1}^{k} (u \mid v_j)(u \mid v_j) - \sum_{i=1}^{k} (u \mid v_i)(v_i \mid u) + \sum_{i=1}^{k} \sum_{j=1}^{k} (u \mid v_i)(u \mid v_j)(v_i \mid v_j)$$

Obsérvese que

$$\sum_{i=1}^{k} \sum_{j=1}^{k} (u \mid v_i)(u \mid v_j)(v_i \mid v_j) = \sum_{i=1}^{k} (u \mid v_i)(u \mid v_i)$$

pues los vectores $v_1, v_2, \ldots, v_k$ forman un conjunto ortonormal.
Entonces

$$\|w\|^2 = (u \mid u) - \sum_{j=1}^{k} (u \mid v_j)(u \mid v_j) - \sum_{i=1}^{k} (u \mid v_i)(u \mid v_i) + \sum_{i=1}^{k} (u \mid v_i)(u \mid v_i)$$

$$= \|u\|^2 - \sum_{j=1}^{k} (u \mid v_j)^2$$

Como, $\|w\|^2 \geq 0$, se obtiene

$$\|u\|^2 - \sum_{j=1}^{k} (u \mid v_j)^2 \geq 0$$

o sea,

$$\|u\|^2 \geq \sum_{j=1}^{k} (u \mid v_j)^2$$

como se quería demostrar.

Por último, en vista del teorema 3.3, el vector u pertenece al espacio generado por el conjunto ortonormal $S = \{v_1, v_2, \ldots, v_k\}$ si, y sólo si

$$w = u - \sum_{j=1}^{k} (u \mid v_j)v_j = 0$$

En tal caso

$$0 = \|w\|^2 = \|u\|^2 - \sum_{j=1}^{k} (u \mid v_j)^2$$

o sea,

$$\|u\|^2 = \sum_{j=1}^{k} (u \mid v_j)^2$$

que es la igualdad de la desigualdad de Bessel.

Q.E.D.

EJERCICIOS (SECCIÓN 3, CAPÍTULO 5)

1. Considere el espacio vectorial $\mathbf{R}^2$ con el producto interno canónico. Encuentre el ángulo entre cada uno de los siguientes pares de vectores:

a) (2, 3) y (2, -1)
b) (1, 2) y (1, 1)
c) (2, 1) y (-3, 2)
d) (1, 0) y (0, 1)
e) (2, 4) y (-6, -3)
f) (2, 3) y (-2, -3)

2. En el espacio $M_{2\times 2}$ con el producto interno canónico, encuentre el ángulo entre los siguientes vectores:

a) $\begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}$ y $\begin{bmatrix} 2 & 4 \\ 2 & 2 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$ y $\begin{bmatrix} 2 & -3 \\ 1 & 4 \end{bmatrix}$

b) $\begin{bmatrix} -2 & 0 \\ 3 & 1 \end{bmatrix}$ y $\begin{bmatrix} 4 & 5 \\ 2 & 2 \end{bmatrix}$

3. En el espacio $C([-1, 1])$ con el producto interno

$$(f \mid g) = \int_{-1}^{1} f(x)g(x)dx$$

encuentre el ángulo entre los siguientes vectores:

a) x y sen x
b) sen x y cos x
c) e^x y x

4. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$, ¿qué ángulo forma un vector $v \in V$ no nulo con su inverso aditivo?

5. En el espacio vectorial $\mathbf{R}^3$ con el producto interno canónico, sean $v = (1, 3, 1)$, $u = (1, 2, -2)$, $w = (0, 0, 1)$, calcule la longitud de los lados y los ángulos internos del triángulo cuyos vértices son los puntos v, u y w.

6. ¿Cuáles de los pares de vectores en los ejercicios 1, 2 y 3 son ortogonales?

7. En el espacio vectorial $\mathbf{R}^3$ considere los vectores

$$v = (2, 3, -1), \quad u = (4, -2, 2)$$

a) Demuestre que si en $\mathbf{R}^3$ se considera el producto interno canónico, los vectores v y u son ortogonales.

b) Calcule el ángulo entre v y u si en $\mathbf{R}^3$ se considera el producto interno

$$(x \mid y) = x_1y_1 + 3x_2y_2 + 8x_3y_3$$

8. En esta sección se observó que el vector cero es un vector ortogonal a todo vector v en un espacio vectorial V con producto interno $(\cdot \mid \cdot)$. Demuestre que éste es el único vector con tal propiedad.

9. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$. Sean v y u vectores en V tales que $\|v\| = \|u\|$, compruebe que los vectores $v + u$ y $v - u$ son ortogonales. Interprete geométricamente este resultado en $\mathbf{R}^2$ con el producto interno canónico.

10. Sea V un espacio vectorial de dimensión finita con el producto interno $(\cdot \mid \cdot)$, demuestre que no existe conjunto ortonormal en V con más de *dim* V elementos.

11. Pruebe que el conjunto de vectores en $\mathbf{R}^4$

$$\{(2, 3, -3, 2), (2, -1, -1, -2) \text{ y } (1, 2, 2, -1)\}$$

es un conjunto ortogonal (con el producto interno canónico en $\mathbf{R}^4$). ¿Es éste un conjunto ortonormal?

12. En el espacio vectorial $C([-1, 1])$ con el producto interno

$$(f \mid g) = \int_{-1}^{1} f(x)g(x)dx$$

demuestre que el conjunto infinito de vectores $S = \{\text{sen } 2n\pi x,\ n = 0, 1, \ldots\}$ es un conjunto ortonormal.

13. Demuestre que el conjunto de vectores en $\mathbf{R}^3$

$$S = \left\{\left(\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right), \left(\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}\right)\right\}$$

es un conjunto ortonormal (con el producto interno canónico de $\mathbf{R}^3$). Use el teorema 3.3 para decidir si cada uno de los siguientes vectores de $\mathbf{R}^3$ pertenecen al subespacio de $\mathbf{R}^3$ generado por S.

a) (1, 2, 1)
b) (4, 5, -2)
c) (3, 3, 0)

14. Compruebe que el conjunto de vectores en $\mathbf{R}^4$

$$S = \left\{\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0\right); \left(0, -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)\right\}$$

es un conjunto ortonormal (con el producto interno canónico de $\mathbf{R}^4$).

Determine los valores de C_1, C_2 y C_3, tales que

$$(x_1, x_2, x_3, x_4) = C_1\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0\right) + C_2\left(0, -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) + C_3\left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$$

② 15. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$. Sean u y v dos vectores de V. Al vector

$$w = \frac{(u \mid v)}{\|v\|^2}v$$

se le llama *proyección ortogonal de u sobre v.*

a) Interprete geométricamente esta definición en el caso del espacio $\mathbf{R}^2$ con el producto interno canónico.

b) Si v es un vector unitario en V, demuestre que la norma de la proyección ortogonal de u sobre v es $|(u|v)|$.

c) Si u es ortogonal a v, ¿cuál es la proyección ortogonal de u sobre v?

d) Reenuncie el teorema de Pitágoras (teorema 3.4) en términos de proyecciones ortogonales del vector u sobre los vectores $v_1, v_2, \ldots, v_k$ del conjunto ortonormal S.

e) Sea $\{v_1, v_2\}$ un conjunto ortonormal de $\mathbf{R}^2$, compruebe que todo vector $u \in \mathbf{R}^2$ se puede escribir de manera única como

$$u = w_1 + w_2$$

en donde w_i es la proyección ortogonal del vector u sobre el vector v_i, $i = 1, 2$. Interprete geométricamente.

f) Más generalmente, si $\{v_1, v_2, \ldots, v_n\}$ es un conjunto ortonormal del espacio vectorial V, entonces todo vector $u \in \mathcal{L}(v_1, v_2, \ldots, v_n)$ se puede escribir de manera única como

$$u = w_1 + w_2 + \ldots + w_n$$

en donde w_i es la proyección ortogonal del vector u sobre el vector v_i, $i = 1, 2, \ldots, n$. Concilie este resultado con el teorema 3.3.

16. Sea $S = \{v_1, v_2, \ldots, v_k\}$ un conjunto ortogonal de vectores en el espacio vectorial V con el producto interno $(\cdot \mid \cdot)$.

Sea

$$u = v_1 + v_2 + \ldots + v_k$$

Demuestre que

$$\|u\|^2 = \|v_1\|^2 + \|v_2\|^2 + \ldots + \|v_n\|^2$$

Concilie este resultado con el teorema de Pitágoras.

③ 17. En este ejercicio se consideran algunos aspectos de la geometría de un cubo n-dimensional en $\mathbf{R}^n$.

Definición. Considere el espacio vectorial $\mathbf{R}^n$ con el producto interno canónico. Sea $\{e_1, e_2, \ldots, e_n\}$ la base canónica de $\mathbf{R}^n$. Se llama "cubo (unitario) de n dimensiones en $\mathbf{R}^n$" al conjunto

$$C = \left\{ \sum_{i=1}^{n} \lambda_i e_i \;\middle|\; 0 \le \lambda_i \le 1,\ i = 1, 2, \ldots, n \right\}$$

A los vectores $e_1, e_2, \ldots, e_n$ se les llama "lados del cubo C".

A un vector de $\mathbf{R}^n$ de la forma

$$d = \sum_{i=1}^{n} \alpha_i e_i$$

en donde $\alpha_i = \pm 1$, $i = 1, 2, \ldots, n$ se le llama "diagonal del cubo C" (se identifica como siendo la misma diagonal del cubo C a los vectores d y $-d$).

a) Dibuje un cubo de una dimensión en $\mathbf{R}$, así como sus diagonales.

b) Dibuje un cubo de dos dimensiones en $\mathbf{R}^2$, así como sus diagonales.

c) Dibuje un cubo de tres dimensiones en $\mathbf{R}^3$, así como sus diagonales.

d) ¿Cuántas diagonales distintas posee un cubo C de n dimensiones en $\mathbf{R}^n$?

e) Sea d una diagonal del cubo C de n dimensiones en $\mathbf{R}^n$. ¿Cuántas diagonales de C son ortogonales a d? Ilustre el resultado en los casos $n = 2$ y $n = 3$.

f) Demuestre que todas las diagonales de un cubo C de n-dimensiones en $\mathbf{R}^n$ tienen la misma longitud. Calcule tal longitud.

g) Compruebe que el ángulo que forma una diagonal d del cubo C de n dimensiones en $\mathbf{R}^n$ con cada uno de sus lados C_i, $i = 1, 2, \ldots, n$ es

$$\varphi = \arccos \frac{1}{\sqrt{n}}$$

h) ¿Para qué valor de n la diagonal d del cubo C de n-dimensiones en $\mathbf{R}^n$ forma con sus lados un ángulo de

h.1) $\pi/4$, h.2) $\pi/3$

i) Sea E la esfera que circunscribe al cubo C de n-dimensiones en $\mathbf{R}^n$. Denote por R al radio de E. Demuestre que para $n = 4$ se tiene $R = 1$ = longitud del lado del cubo C, mientras que para $n > 4$, $R > 1$. ¿Qué ocurre con los casos $n = 2, 3$? Haga un dibujo en cada uno de estos casos.

j) Pruebe que la norma de la proyección ortogonal (véase ejercicio 15) de un lado cualquiera del cubo C de n-dimensiones en $\mathbf{R}^n$ sobre una diagonal d cualquiera de C, es la n-ésima parte de la longitud de la diagonal d. Verifique este resultado geométricamente en los casos $n = 2$ y $n = 3$.

4. BASES ORTONORMALES

Sea $(V, (\cdot \mid \cdot))$ un espacio de dimensión finita con producto interno, se dice que la base $\beta = \{v_1, v_2, \ldots, v_n\}$ de V es una *base ortonormal* si el conjunto β es un conjunto ortonormal de vectores en V.

Obsérvese que el conjunto ortonormal de vectores en V, $\beta = \{v_1, v_2, \ldots, v_k\}$ es una base ortonormal de V si, y sólo si este conjunto genera a V. (La parte "sólo si" de esta afirmación es obvia. La parte "si" se sigue del hecho de que el conjunto β es linealmente independiente —corolario del teorema 3.2).

EJEMPLO 1 El ejemplo más simple de bases ortonormales se encuentra en la base canónica de $\mathbf{R}^n$ (con el producto interno canónico).

En efecto, la base $\beta = \{e_1, e_2, \ldots, e_n\}$ de $\mathbf{R}^n$ en donde

$$e_i = (0, \ldots, \underbrace{1}_{i}, \ldots, 0) \in \mathbf{R}^n$$

es una base ortonormal pues

$$(e_i \mid e_j) = 0 \quad \text{si} \quad i \neq j$$
$$\|e_i\| = 1 \quad i = 1, 2, \ldots, n$$

TEOREMA 4.1 Sea $\beta = \{v_1, v_2, \ldots, v_k\}$ un conjunto ortonormal de vectores en el espacio con producto interno V. Las siguientes afirmaciones son equivalentes:

1) β es una base de V.
2) Para cualquier vector $u \in V$ se tiene

$$u = \sum_{i=1}^{k} (u \mid v_i) v_i$$

3) Para cualesquiera dos vectores $u, w \in V$ se tiene

$$(u \mid w) = \sum_{i=1}^{k} (u \mid v_i)(v_i \mid w)$$

4) Para cualquier vector $u \in V$ se tiene

$$\|u\|^2 = \sum_{i=1}^{k} (u \mid v_i)^2$$

DEMOSTRACIÓN Se probará 1) $\Rightarrow$ 2) $\Rightarrow$ 3) $\Rightarrow$ 4) $\Rightarrow$ 1)

1) $\Rightarrow$ 2). Si β es una base de V, entonces β genera a V. Por tanto, el vector u se encuentra en el espacio generado por el conjunto ortonormal V. 2) se concluye entonces del teorema 3.3.

2) $\Rightarrow$ 3). Sean u y w dos vectores de V. Se escribe

$$u = \sum_{i=1}^{k} (u \mid v_i)v_i$$

Entonces

$$(u \mid w) = \sum_{i=1}^{k} ((u \mid v_i)v_i \mid w) = \sum_{i=1}^{k} (u \mid v_i)(v_i \mid w)$$

lo que prueba 3).

3) $\Rightarrow$ 4). Tómese en 3) $w = u$. Entonces

$$\|u\|^2 = (u \mid u) = \sum_{i=1}^{k} (u \mid v_i)(v_i \mid u) = \sum_{i=1}^{k} (u \mid v_i)^2$$

4) $\Rightarrow$ 1). Se tiene por hipótesis que para cualquier vector $u \in V$ se obtiene

$$\|u\|^2 = \sum_{i=1}^{k} (u \mid v_i)^2$$

Ésta es precisamente la igualdad de la desigualdad de Bessel (teorema 3.6) la cual se sabe que se cumple si, y sólo si el vector u se encuentra en el espacio generado por $v_1, v_2, \ldots, v_n$. Como u es un vector arbitrario de V, el conjunto $\beta = \{v_1, v_2, \ldots, v_n\}$ debe generar a todo el espacio V. Entonces por la observación previa al teorema, β es una base de V.

Q.E.D.

EJEMPLO 2 El hecho de disponer de una base ortonormal en un espacio de dimensión finita con producto interno tiene, según el teorema, ventajas de carácter "práctico". Véase un ejemplo.

La base

$$\beta = \{v_1, v_2, v_3\} = \left\{\left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right), \left(\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right), \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)\right\} \text{ de } R^3$$

(con el producto interno canónico) es una base ortonormal, pues

$$(v_1 \mid v_2) = \left(\frac{2}{3}\right)\left(\frac{2}{3}\right) + \left(-\frac{2}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)\left(-\frac{2}{3}\right) = 0$$

$$(v_1 \mid v_3) = \left(\frac{2}{3}\right)\left(\frac{1}{3}\right) + \left(-\frac{2}{3}\right)\left(\frac{2}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = 0$$

$$(v_2 \mid v_3) = \left(\frac{2}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{2}{3}\right) + \left(-\frac{2}{3}\right)\left(\frac{2}{3}\right) = 0$$

y

$$\|v_1\|^2 = \left(\frac{2}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 = 1$$

$$\|v_2\|^2 = \left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 = 1$$

$$\|v_3\|^2 = \left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2 = 1$$

Supóngase que se quiere expresar el vector $u = (6, 6, -9)$ como una combinación lineal de los vectores de la base β. En principio, esto conduciría a resolver un sistema de 3 ecuaciones con 3 incógnitas. Sin embargo, el teorema 4.1 (la afirmación 2) permite escribir

$$\begin{aligned} u &= (u \mid v_1)v_1 + (u \mid v_2)v_2 + (u \mid v_3)v_3 \\ &= \left((6)\left(\frac{2}{3}\right) + (6)\left(-\frac{2}{3}\right) + (-9)\left(\frac{1}{3}\right)\right)v_1 + \left((6)\left(\frac{2}{3}\right) + (6)\left(\frac{1}{3}\right) + (-9)\left(-\frac{2}{3}\right)\right)v_2 \\ &\quad + \left((6)\left(\frac{1}{3}\right) + (6)\left(\frac{2}{3}\right) + (-9)\left(\frac{2}{3}\right)\right)v_3 \\ &= -3v_1 + 12v_2 \end{aligned}$$

El próximo objetivo es mostrar que todo espacio vectorial V de dimensión finita con producto interno posee una base ortonormal.

4.1. EL PROCESO DE GRAM-SCHMIDT

Sea $(V, (\cdot \mid \cdot))$ un espacio vectorial de dimensión finita con producto interno. Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de él.

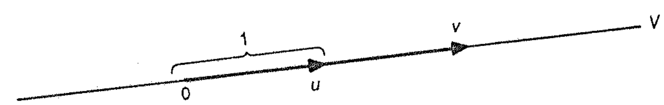
Se va a demostrar que el espacio V tiene una base ortonormal. Para lograr esto, *se exhibirá* tal base, usando para ello un procedimiento que permite construir, a

partir de la base β, una base ortonormal. Este procedimiento es conocido como "proceso (de ortonormalización de bases) de Gram-Schmidt".

Se analizarán primeramente los casos $n = 1, 2$ y 3. Si $n = 1$ (el espacio V tiene dimensión 1) y $\beta = \{v_1\}$ es una base de V, para pasar de esta base a una base ortonormal sólo se tiene que ajustar la norma de v_1 al valor 1. Sea

$$u_1 = \frac{v_1}{\|v_1\|}$$

Entonces es claro que: 1) $\mathcal{L}(u_1) = \mathcal{L}(v_1) = V$, 2) $\|u_1\| = 1$, de modo que $\beta_{ON} = \{u_1\}$ es la base ortonormal deseada



Véase ahora el caso $n = 2$.

Se tiene entonces la base $\beta = \{v_1, v_2\}$ de V y se quiere construir a partir de ella una base ortonormal $\beta_{ON} = \{u_1, u_2\}$. Para construir u_1, sólo se ajusta la norma de v_1 al valor 1 (como en el caso anterior), es decir, *se normaliza* el vector v_1.

$$u_1 = \frac{v_1}{\|v_1\|}$$

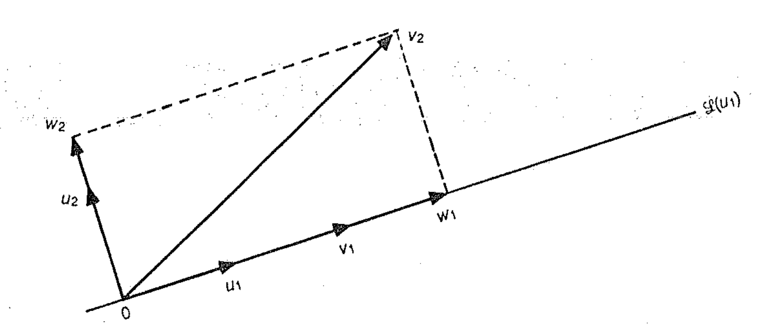
La idea para construir el vector u_2 es *descomponer* el vector v_2 como una suma de dos vectores $w_1 + w_2$, en donde w_1 es un vector que se encuentra en el espacio que genera u_1 (el vector previamente construido) y w_2 es un vector que es ortogonal a u_1. Es decir,

$$v_2 = w_1 + w_2$$

en donde

$$w_1 \in \mathcal{L}(u_1) \quad \text{y} \quad (w_2 \mid u_1) = 0$$

Geométricamente esto se vería como



En tal caso, el vector u_2 procurado sería simplemente

$$u_2 = \frac{w_2}{\|w_2\|}$$

Encuéntrese pues el vector w_2.

Como $w_1 \in \mathcal{L}(u_1)$ se puede escribir

$$w_1 = cu_1$$

para algún escalar $c \in \mathbf{R}$. [El valor de c se halla imponiendo la condición $(w_2 \mid u_1) = 0$.]. Entonces

$$w_2 = v_2 - w_1 = v_2 - cu_1$$

Se quiere que $(w_2 \mid u_1) = 0$, es decir,

$$\begin{aligned} 0 = (w_2 \mid u_1) &= (v_2 - cu_1 \mid u_1) = (v_2 \mid u_1) - c(u_1 \mid u_1) = (v_2 \mid u_1) - c\,\|u_1\|^2 \\ &= (v_2 \mid u_1) - c \end{aligned}$$

de donde se ve que el valor de c debe ser

$$c = (v_2 \mid u_1)$$

y entonces el vector w_2 sería

$$w_2 = v_2 - cu_1 = v_2 - (v_2 \mid u_1)u_1$$

y finalmente

$$u_2 = \frac{v_2 - (v_2 \mid u_1)u_1}{\|v_2 - (v_2 \mid u_1)u_1\|}$$

Se comprueba que en efecto $\beta_{ON} = \{u_1, u_2\}$ es una base ortonormal de V.

Véase por último el caso $n = 3$. A partir de la base $\beta = \{v_1, v_2, v_3\}$ de V se quiere construir una base ortonormal $\beta_{ON} = \{u_1, u_2, u_3\}$.

Se escribe

$$u_1 = \frac{v_1}{\|v_1\|} \tag{4.1}$$

El vector u_2 se halla como en el caso anterior

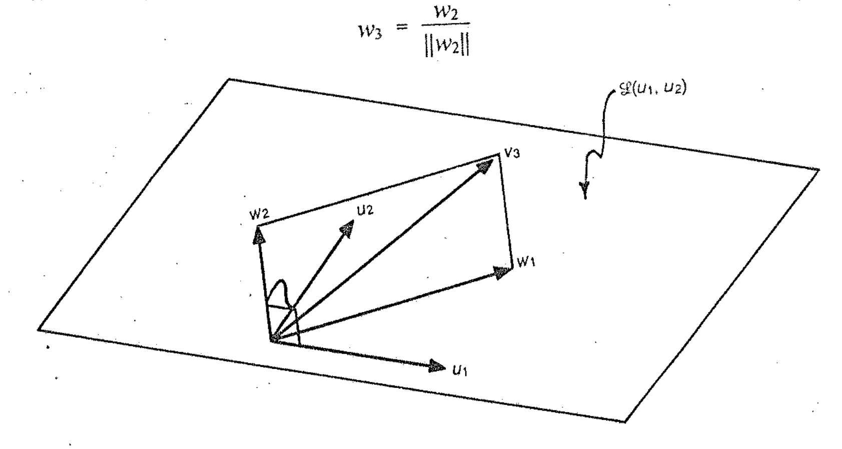
$$u_2 = \frac{v_2 - (v_2 \mid u_1)u_1}{\|v_2 - (v_2 \mid u_1)u_1\|} \tag{4.2}$$

Para construir el vector u_3 (a partir de v_3 y los vectores previamente construidos u_1 y u_2) se usa la misma idea que en el caso anterior: se presenta el vector v_3 como

una suma $w_1 + w_2$, en donde w_1 es un vector que se encuentra en el espacio generado por u_1 y u_2, y w_2 es ortogonal a u_1 y a u_2. (Como se muestra en la figura.)

En tal caso, el vector u_3 procurado sería simplemente

$$w_3 = \frac{w_2}{\|w_2\|}$$



Escríbase entonces

$$v_3 = w_1 + w_2$$

en donde $w_1 \in \mathcal{L}(u_1, u_2)$ y $(w_2 \mid u_1) = (w_2 \mid u_2) = 0$ y hállese el vector w_2.

Como $w_1 \in \mathcal{L}(u_1, u_2)$ existen escalares c y d tales que

$$w_1 = cu_1 + du_2$$

Los valores de estos escalares se hallan imponiendo las condiciones $(w_2 \mid u_1) = 0$ y $(w_2 \mid u_2) = 0$.

Como

$$w_2 = v_3 - w_1 = v_3 - cu_1 - du_2$$

se tiene

$$0 = (w_2 \mid u_1) = (v_3 - cu_1 - du_2 \mid u_1) = (v_3 \mid u_1) - c\underbrace{(u_1 \mid u_1)}_{\|u_1\|^2 = 1} - d\underbrace{(u_2 \mid u_1)}_{= 0} = (v_3 \mid u_1) - c$$

y

$$0 = (w_2 \mid u_2) = (v_3 - cu_1 - du_2 \mid u_2) = (v_3 \mid u_2) - c\underbrace{(u_1 \mid u_2)}_{= 0} - d\underbrace{(u_2 \mid u_2)}_{\|u_1\|^2 = 1} = (v_3 \mid u_2) - d$$

de donde

$$c = (v_3 \mid u_1) \qquad d = (v_3 \mid u_2)$$

y entonces

$$w_2 = v - w_1 = v_3 - (v_3 \mid u_1)u_1 - (v_3 \mid u_2)u_2$$

por lo que el vector u_3 procurado es:

$$u_3 = \frac{v_3 - (v_3 \mid u_1)u_1 - (v_3 \mid u_2)u_2}{\|v_3 - (v_3 \mid u_1)u_1 - (v_3 \mid u_2)u_2\|} \tag{4.3}$$

Se tiene en efecto que $(u_3 \mid u_1) = (u_3 \mid u_2) = 0$, $\|u_3\| = 1$, por lo que $\beta_{ON} = \{u_1, u_2, u_3\}$ es una base ortonormal de V.

Con el análisis de estos 3 casos particulares se puede generalizar el procedimiento para construir una base ortonormal β_{ON} a partir de una base dada β del espacio V de dimensión n.

De hecho no es difícil "adivinar", a partir de las fórmulas (4.1), (4.2) y (4.3), que definen a u_1, u_2, y u_3, respectivamente, que el vector u_4 sería

$$u_4 = \frac{u_4 - (v_4 \mid u_1)u_1 - (v_4 \mid u_2)u_2 - (v_4 \mid u_3)u_3}{\|u_4 - (v_4 \mid u_1)u_1 - (v_4 \mid u_2)u_2 - (v_4 \mid u_3)u_3\|}$$

Más generalmente, supóngase que ya se han construido los primeros k vectores $u_1, u_2, \ldots, u_k$ $(1 \leq k < n)$ de la base ortonormal. Entonces el vector u_{k+1} se construirá a partir de ellos y del vector v_{k+1} de la base dada, siendo

$$u_{k+1} = \frac{v_{k+1} - (v_{k+1} \| u_1)u_1 - (v_{k+1} \mid u_2)u_2 - \ldots - (v_{k+1} \mid u_k)u_k}{\|v_{k+1} - (v_{k+1} \mid u_1)u_1 - (v_{k+1} \mid u_2)u_2 - \ldots - (v_{k+1} \mid u_k)u_k\|}$$

La discusión anterior se resume en el siguiente teorema:

TEOREMA 4.2

Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base del espacio V con producto interno $(\cdot \mid \cdot)$. Entonces el conjunto $\beta_{ON} = \{u_1, u_2, \ldots, u_n\}$ en donde los vectores $u_1, u_2, \ldots, u_n$ son definidos inductivamente como

$$u_1 = \frac{v_1}{\|v_1\|}$$

$$u_2 = \frac{v_2 - (v_2 \mid u_1)u_1}{\|v_2 - (v_2 \mid u_1)u_1\|}$$

$$u_3 = \frac{v_3 - (v_3 \mid u_1)u_1 - (v_3 \mid u_2)u_2}{\|v_3 - (v_3 \mid u_1)u_1 - (v_3 \mid u_2)u_2\|}$$

$$\vdots$$

$$u_k = \frac{v_k - \sum_{i=1}^{k-1} (v_k \mid u_i)u_i}{\|v_k - \sum_{i=1}^{k-1} (v_k \mid u_i)u_i\|}$$

$$\vdots$$

$$u_n = \frac{v_n - \sum_{i=1}^{n-1} (v_n \mid u_i)u_i}{\|v_n - \sum_{i=1}^{n-1} (v_n \mid u_i)u_i\|}$$

forma una base ortonormal de V.

DEMOSTRACIÓN Al proceder por inducción sobre n supóngase válido el teorema para $n = k$ (el caso $n = 1$ es claro) y pruébese su validez para $n = k + 1$.

Sea entonces $\beta = \{v_1, v_2, \ldots, v_k, v_{k+1}\}$ una base del espacio V de dimensión $k + 1$.

Por hipótesis de inducción, el subespacio W de V generado por $\{v_1, v_2, \ldots, v_k\}$ tiene como base ortonormal a los vectores $u_1, u_2, \ldots, u_k$.

Sea

$$u_{k+1} = \frac{v_{k+1} - \sum_{i=1}^{k} (v_{k+1} \mid u_i)u_i}{\|v_{k+1} - \sum_{i=1}^{k} (v_{k+1} \mid u_i)u_i\|}$$

Ciertamente $u_{k+1} \neq 0$, pues caso contrario se tendría

$$v_{k+1} = \sum_{i=1}^{k} (v_{k+1} \mid u_i)u_i$$

lo que implica que

$$v_{k+1} \in \mathscr{L}(u_1, u_2, \ldots, u_k) = \mathscr{L}(v_1, v_2, \ldots, v_k)$$

lo cual contradice el hecho de que $v_1, v_2, \ldots, v_{k+1}$ es un conjunto linealmente independiente (pues es una base de V).

Se afirma que $(u_{k+1} \mid u_j) = 0, j = 1, 2, \ldots, k$.

Para simplificar la notación llámese

$$N = \|v_{k+1} - \sum_{i=1}^{k} (v_{k+1} \mid u_i)u_i\|$$

Entonces

$$(u_{k+1} \mid u_j) = \frac{1}{N}\left(v_{k+1} - \sum_{i=1}^{k} (v_{k+1} \mid u_i)u_i \mid u_j\right)$$

$$= \frac{1}{N}(v_{k+1} \mid u_j) - \frac{1}{N}\sum_{i=1}^{k} (v_{k+1} \mid u_i)(u_i \mid u_j)$$

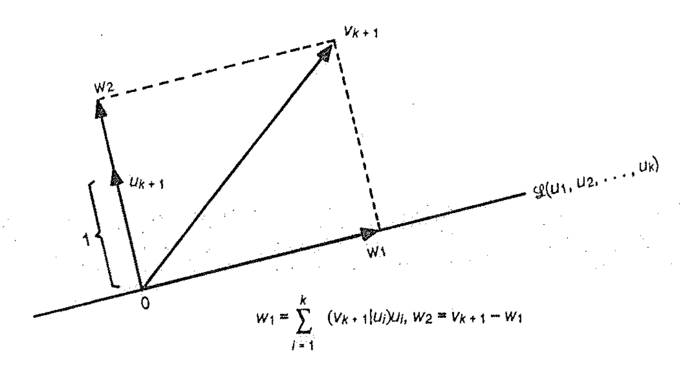
$$= \frac{1}{N}(v_{k+1} \mid u_j) - \frac{1}{N}(v_{k+1} \mid u_j) = 0$$

como se quería probar. Además es claro que $\|u_{k+1}\| = 1$.

Como los vectores $u_1, u_2, \ldots, u_{k+1}$ forman un conjunto ortonormal de $k + 1$ vectores en el espacio V de dimensión $k + 1$, éstos deben ser una base (ortonormal) de V.

Q.E.D.

Según el análisis presentado en los casos concretos $\dim V = 1, 2$ y 3, se podría representar esquemáticamente la construcción del vector u_{k+1} como



Un corolario obvio del teorema anterior es:

COROLARIO Todo espacio vectorial de dimensión finita con producto interno posee una base ortonormal.

DEMOSTRACIÓN Sea β una base cualquiera del espacio dado, aplique el teorema anterior para obtener, a partir de β, una base ortonormal.

Q.E.D.

Véanse un par de ejemplos.

EJEMPLO 3 Considere en el espacio $\mathbf{R}^3$ con el producto interno canónico la base

$$\beta = \{v_1, v_2, v_3\} = \{(1, 1, 0), (1, 1, 1), (0, 1, 1)\}$$

A partir de β obtenga una base ortonormal de $\mathbf{R}^3$.

El primer vector u_1 de la base β_{ON} procurada es simplemente

$$u_1 = \frac{1}{\|v_1\|}v_1 = \frac{1}{\sqrt{(1)^2 + (1)^2 + (0)^2}}(1, 1, 0) = \frac{1}{\sqrt{2}}(1, 1, 0)$$

Para hallar u_2 calcule

$$v_2 - (v_2 \mid u_1)u_1 = (1, 1, 1) - \frac{1}{\sqrt{2}}[(1)(1) + (1)(1) + (1)(0)]\frac{1}{\sqrt{2}}(1, 1, 0)$$

$$= (1, 1, 1) - \frac{1}{2}(2)(1, 1, 0) = (1, 1, 1) - (1, 1, 0) = (0, 0, 1)$$

Como la norma de este vector es 1, el vector u_2 procurado es:

$$u_2 = (0, 0, 1)$$

Por último, para hallar u_3 se calcula

$$v_3 - (v_3 \mid u_1)u_1 - (v_3 \mid u_2)u_2 = (0, 1, 1) - \frac{1}{\sqrt{2}}[(0)(1) + (1)(1) + (1)(0)]\frac{1}{\sqrt{2}}(1, 1, 0)$$

$$- [(0)(0) + (1)(0) + (1)(1)]\,(0, 0, 1)$$

$$= (0, 1, 1) - \frac{1}{2}(1, 1, 0) - (0, 0, 1) = \left(-\frac{1}{2}, \frac{1}{2}, 0\right)$$

Entonces

$$u_3 = \frac{1}{\left\|\left(-\frac{1}{2}, \frac{1}{2}, 0\right)\right\|}\left(-\frac{1}{2}, \frac{1}{2}, 0\right) = \sqrt{2}\left(-\frac{1}{2}, \frac{1}{2}, 0\right)$$

y así,

$$\beta_{ON} = \left\{\frac{1}{\sqrt{2}}(1, 1, 0), (0, 0, 1), \sqrt{2}\left(-\frac{1}{2}, \frac{1}{2}, 0\right)\right\}$$

es una base ortonormal de $\mathbf{R}^3$.

EJEMPLO 4 Considere ahora el espacio vectorial P_2 con el producto interno

$$(p \mid q) = \int_0^1 p(x)q(x)dx$$

Transforme la base canónica de P_2, $\beta = \{1, x, x^2\}$ en una base ortonormal.

Como

$$\|v_1\| = \sqrt{\int_0^1 (1)^2 dx} = 1$$

el vector u_1 de la base ortonormal es $u_1 = 1$.

Para hallar u_2, calcule

$$v_2 - (v_2 \mid u_1)u_1 = x - \left(\int_0^1 x\,dx\right)(1) = x - \frac{1}{2}$$

La norma de este vector es:

$$\left\|x - \frac{1}{2}\right\| = \sqrt{\int_0^1 \left(x - \frac{1}{2}\right)^2 dx} = \frac{1}{\sqrt{12}}$$

por lo que el vector u_2 será

$$u_2 = \sqrt{12}\left(x - \frac{1}{2}\right)$$

Para hallar u_3 calcule

$$v_3 - (v_3 \mid u_1)u_1 - (v_3 \mid u_2)u_2 = x^2 - \left(\int_0^1 x^2\,dx\right)(1) - \sqrt{12}\left(\int_0^1 x^2\left(x - \frac{1}{2}\right)dx\right)$$

$$\cdot \sqrt{12} \cdot \left(x - \frac{1}{2}\right)$$

$$= x^2 - \frac{1}{3} - \left(x - \frac{1}{2}\right) = x^2 - x + \frac{1}{6}$$

Como

$$\left\|x^2 - x + \frac{1}{6}\right\| = \sqrt{\int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2 dx} = \frac{1}{\sqrt{180}}$$

se tiene que

$$u_3 = \sqrt{180}\left(x^2 - x + \frac{1}{6}\right)$$

de modo que

$$\beta_{ON} = \left\{1, \sqrt{12}\left(x - \frac{1}{2}\right), \sqrt{180}\left(x^2 - x + \frac{1}{6}\right)\right\}$$

es una base ortonormal de P_2.

4.2. MATRICES ORTOGONALES

Considérese el espacio $\mathbf{R}^3$ con el producto interno canónico. Ciertamente la base

$$\beta_1 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

es una base ortonormal de $\mathbf{R}^3$.

En la sección anterior se vio que

$$\beta_2 = \left\{\frac{1}{\sqrt{2}}(1, 1, 0), (0, 0, 1), \sqrt{2}\left(-\frac{1}{2}, \frac{1}{2}, 0\right)\right\}$$

era también una base ortonormal de $\mathbf{R}^3$.

La matriz P de cambio de base de β_2 a β_1 se obtiene, como se sabe, expresando cada vector de la base β_2 en términos de la base β_1 y colocando los vectores de coordenadas correspondientes como columnas de la matriz P.

En este caso se tiene

$$P = \begin{bmatrix} 1/\sqrt{2} & 0 & -\sqrt{2}/2 \\ 1/\sqrt{2} & 0 & \sqrt{2}/2 \\ 0 & 1 & 0 \end{bmatrix}$$

Obsérvese que

$$P^tP = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 0 & -\sqrt{2}/2 \\ 1/\sqrt{2} & 0 & \sqrt{2}/2 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

de donde

$$P^{-1} = P^t$$

En realidad esto es un hecho general que se enuncia y prueba a continuación.

TEOREMA 4.3

Sea $(V, (\cdot \mid \cdot))$ un espacio vectorial de dimensión finita con producto interno.
Sean $\beta_1 = \{v_1, v_2, \ldots, v_n\}$, $\beta_2 = \{u_1, u_2, \ldots, u_n\}$ dos bases ortonormales de V. Si P es la matriz (de orden n) de cambio de base de β_2 a β_1, entonces (P es inversible y)

$$P^{-1} = P^t$$

DEMOSTRACIÓN Como β_1 es una base ortonormal de V, el vector u_j, $1 \le j \le n$, queda expresado como combinación lineal de la base β_1 según la fórmula

$$u_j = (u_j \mid v_1)v_1 + (u_j \mid v_2)v_2 + \ldots + (u_j \mid v_n)v_n$$

[teorema 4.1 (2)].

Entonces la matriz P tiene por elementos en su j-ésima columna a $(u_j \mid v_1)$, $(u_j \mid v_2), \ldots, (u_j \mid v_n)$.

Es decir,

$$P = \begin{bmatrix} (u_1 \mid v_1) & (u_2 \mid v_1) & \ldots & (u_n \mid v_1) \\ (u_1 \mid v_2) & (u_2 \mid v_2) & \ldots & (u_n \mid v_2) \\ & \cdot\ \cdot\ \cdot & & \\ (u_1 \mid v_n) & (u_2 \mid v_n) & \ldots & (u_n \mid v_n) \end{bmatrix}$$

Hágase el producto P^tP.

$$P^tP = \begin{bmatrix} (u_1 \mid v_1) & (u_1 \mid v_2) & \ldots & (u_1 \mid v_n) \\ (u_2 \mid v_1) & (u_2 \mid v_2) & \ldots & (u_2 \mid v_n) \\ & \cdot\ \cdot\ \cdot & & \\ (u_n \mid v_1) & (u_n \mid v_2) & \ldots & (u_n \mid v_n) \end{bmatrix}\begin{bmatrix} (u_1 \mid v_1) & (u_2 \mid v_1) & \ldots & (u_n \mid v_1) \\ (u_1 \mid v_2) & (u_2 \mid v_2) & \ldots & (u_n \mid v_2) \\ & \cdot\ \cdot\ \cdot & & \\ (u_1 \mid v_n) & (u_2 \mid v_n) & \ldots & (u_n \mid v_n) \end{bmatrix}$$

llámese c_{ij} al elemento de la i-ésima línea y la j-ésima columna de P^tP. Entonces

$$c_{ij} = (u_i \mid v_1)(u_j \mid v_1) + (u_i \mid v_2)(u_j \mid v_2) + \ldots + (u_i \mid v_n)(u_j \mid v_n) = \sum_{k=1}^{n} (u_i \mid v_k)(u_j \mid v_k)$$

$$= \sum_{k=1}^{n} (u_i \mid v_k)(v_k \mid u_j) = (u_i \mid u_j)$$

en donde en la última igualdad se usa el teorema 4.1, inciso (3).

Entonces, por ser β_2 una base ortonormal se tiene que

$$c_{ij} = (u_i \mid u_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Es decir,

$$P^tP = I_n$$

o bien,

$$P^{-1} = P^t$$

Q.E.D.

DEFINICIÓN Se dice que la matriz inversible A es una matriz *ortogonal* si $A^{-1} = A^t$.

Obsérvese que si A es una matriz ortogonal, su inversa también lo es, pues si $B = A^{-1}$ se tiene $B^{-1} = A = (A^t)^t = (A^{-1})^t = B^t$.

Lo que establece entonces el teorema 4.3 es que las matrices de cambio de una base ortonormal a otra base ortonormal en un espacio vectorial de dimensión finita con producto interno, son matrices ortogonales.

El siguiente teorema establece una propiedad importante de las matrices ortogonales.

TEOREMA 4.4 Si A es una matriz (de orden n) ortogonal, entonces el determinante de A es 1 o −1.

DEMOSTRACIÓN Se sabe que $\det A = \det A^t$. Entonces siendo A ortogonal se tiene $AA^t = I$, de donde: $\det(AA^t) = \det I$, o sea, $(\det A)(\det A^t) = 1$, esto es, $(\det A)^2 = 1$ y por lo tanto, $\det A$ es 1 o −1.

Q.E.D.

En la siguiente sección se estudiarán aquellos operadores lineales en un espacio vectorial de dimensión finita con producto interno que son representados por matrices ortogonales.

Pero antes, para finalizar la presente sección, se explotará la existencia de bases ortonormales en un espacio vectorial V de dimensión finita con producto interno para demostrar que el espacio V puede siempre presentarse como una suma directa de un subespacio dado de él y su "complemento ortogonal".

4.3. COMPLEMENTOS ORTOGONALES

Sea V un espacio con producto interno y sea W un subespacio de él. Al conjunto

$$W^{\perp} = \{v \in V \mid (v \mid w) = 0 \quad \forall\, w \in W\}$$

se le llama el *complemento ortogonal* de W.

Es decir, el complemento ortogonal del subespacio W de V está constituido por aquellos vectores de V que son ortogonales a todos los vectores de W.

Obsérvese que el conjunto $W^{\perp}$ no es vacío, pues al menos contiene al vector $0 \in V$.

En realidad, el conjunto $W^{\perp}$, es también un subespacio de V, como se mostrará a continuación:

Si $v_1, v_2 \in W^{\perp}$, entonces para cualquier $w \in W$ se tiene

$$(v_1 + v_2 \mid w) = (v_1 \mid w) + (v_2 \mid w) = 0 + 0 = 0$$

por lo que $v_1 + v_2 \in W^{\perp}$.

Similarmente, si $c \in \mathbf{R}$ y $v \in W^{\perp}$ se tiene

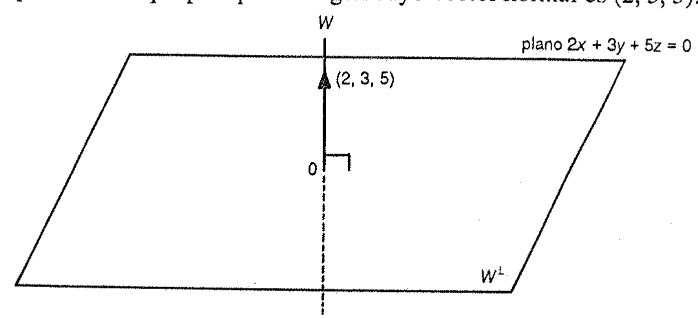
$$(cv \mid w) = c(v \mid w) = c \cdot 0 = 0 \qquad \forall\, w \in W$$

de donde $cv \in W^{\perp}$, y por tanto, $W^{\perp}$ es un subespacio de V.

EJEMPLO 5 Por ejemplo, si se considera en el espacio $\mathbf{R}^3$ con el producto interno canónico, el subespacio W generado por el vector (2, 3, 5), se tiene que

$$\begin{aligned} W^{\perp} &= \{(x, y, z) \in \mathbf{R}^3 \mid ((x, y, z) \mid (x_0, y_0, z_0)) = 0 \quad \forall\,(x_0, y_0, z_0) \in \mathcal{L}((2, 3, 5))\} \\ &= \{(x, y, z) \in \mathbf{R}^3 \mid ((x, y, z) \mid c(2, 3, 5)) = 0 \quad \forall\, c \in \mathbf{R}\} \\ &= \{(x, y, z) \in \mathbf{R}^3 \mid c(2x + 3y + 5z) = 0 \quad \forall\, c \in \mathbf{R}\} \\ &= \{(x, y, z) \in \mathbf{R}^3 \mid 2x + 3y + 5z = 0\} \end{aligned}$$

que es el plano en $\mathbf{R}^3$ que pasa por el origen cuyo vector normal es (2, 3, 5).



El siguiente teorema es el resultado principal de esta subsección.

TEOREMA 4.5 Sea $(V, (\cdot \mid \cdot))$ un espacio vectorial de dimensión finita con producto interno y sea W un subespacio de V. Entonces

$$V = W \oplus W^{\perp}$$

DEMOSTRACIÓN Sea $\beta_1 = \{v_1, v_2, \ldots, v_k\}$ ($k \le \dim V$) una base del subespacio W. Aplicando el proceso de Gram-Schmidt, se puede construir, a partir de β_1, una base ortonormal $\beta_1' = \{u_1, u_2, \ldots, u_k\}$ de W. Se extiende esta base a una base β de V y se aplica nuevamente el proceso de Gram-Schmidt a la base β para obtener una base ortonormal β' de V. Es claro que los primeros k vectores de β' son $u_1, u_2, \ldots, u_k$ (los vectores de la base ortonormal de W). Es decir,

$$\beta' = \{u_1, u_2, \ldots, u_k, u_{k+1}, \ldots, u_n\}$$

en donde $n = \dim V$.

Se verá que los vectores $u_{k+1}, \ldots, u_n$ forman una base (ortonormal) de $W^{\perp}$.

Ciertamente $u_{k+1}, \ldots, u_n$ son vectores linealmente independientes, pues ellos constituyen un conjunto ortonormal de V (corolario del teorema 3.2).

Sea $v \in W^{\perp}$. Como $v \in V$, existen escalares $c_1, c_2, \ldots, c_n$ tales que

$$v = c_1u_1 + c_2u_2 + \ldots + c_ku_k + c_{k+1}u_{k+1} + \ldots + c_nu_n = \sum_{j=1}^{n} c_ju_j$$

Ya que $(v \mid u_i) = 0$ para $i = 1, 2, \ldots, k$, se tiene

$$0 = (v \mid u_i) = \left(\sum_{j=1}^{n} c_j u_j \mid u_i \right) = \sum_{j=1}^{n} c_j (u_j \mid u_i) = c_i$$

Es decir, $c_i = 0$ para $i = 1, 2, \ldots, k$.
Entonces

$$v = c_{k+1}u_{k+1} + \ldots + c_n u_n$$

lo que muestra que los vectores $u_{k+1}, \ldots, u_n$ generan a $W^{\perp}$ y por lo tanto que constituyen una base de él.
Dado entonces cualquier vector $v \in V$ puede expresarse como

$$v = \underbrace{c_1 u_1 + \ldots + c_k u_k}_{\in W} + \underbrace{c_{k+1}u_{k+1} + \ldots + c_n u_n}_{\in W^{\perp}}$$

Es decir,

$$V = W + W^{\perp}$$

Supóngase por último que $w \in W \cap W^{\perp}$, entonces $(w \mid w) = 0$, lo cual implica que $w = 0$. Es decir, que $W \cap W^{\perp} = \{0\}$, y entonces por el teorema 3.6 del capítulo 3 se tiene

$$V = W \oplus W^{\perp}$$

Q.E.D.

De hecho, se puede ser más explícito con la conclusión del teorema anterior.
Sea $\beta = \{u_1, u_2, \ldots, u_k\}$ una base ortonormal del subespacio W. Entonces todo vector $v \in V$ se puede escribir de manera única como

$$v = v_1 + v_2, \qquad v_1 \in W,\ v_2 \in W^{\perp}$$

en donde

$$v_1 = \sum_{i=1}^{k} (v \mid u_i) u_i$$

y

$$v_2 = v - v_1$$

En efecto, ciertamente $v = v_1 + v_2$ y $v_1 \in W$, pues v_1 es una combinación lineal de los vectores de la base β de W.
Sólo falta verificar que $v_2 = v - v_1 \in W^{\perp}$.

Pero, para cualquier $w \in W$ se tiene

$$w = \sum_{j=1}^{k} c_j u_j$$

y

$$(v_2 \mid w) = (v - v_1 \mid w) = \left(v - \sum_{i=1}^{k} (v \mid u_i) u_i \,\middle|\, \sum_{j=1}^{k} c_j u_j \right)$$

$$= \sum_{j=1}^{k} c_j (v \mid u_j) - \sum_{i=1}^{k} \sum_{j=1}^{k} (v \mid u_i) c_j (u_i \mid u_j)$$

$$= \sum_{j=1}^{k} c_j (v \mid u_j) - \sum_{i=1}^{k} c_i (v \mid u_i) = 0$$

Lo que muestra que efectivamente $v_2 \in W^{\perp}$.

EJEMPLO 6

Por ejemplo, considere el espacio $\mathbf{R}^4$ con el producto interno canónico. Sea $W = \mathcal{L}(w_1, w_2)$ en donde

$$w_1 = (1, 1, 0, 0) \qquad w_2 = (0, 1, 0, 1)$$

Suponga que se quiere expresar el vector $v = (2, 3, -1, 4) \in \mathbf{R}^4$ como $v = v_1 + v_2$, en donde $v_1 \in W$ y $v_2 \in W^{\perp}$.
Para aplicar las ideas anteriormente analizadas, se necesita primeramente disponer de una base ortonormal de W. Sea $\beta = \{u_1, u_2\}$ tal base. Se procede entonces con el proceso de Gram-Schmidt a obtener β.

$$u_1 = \frac{1}{\|(1, 1, 0, 0)\|}(1, 1, 0, 0) = \frac{1}{\sqrt{2}}(1, 1, 0, 0)$$

$$v_2 - (v_2 \mid u_1)u_1 = (0, 1, 0, 1) - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}(1, 1, 0, 0) = \left(-\frac{1}{2}, \frac{1}{2}, 0, 1\right)$$

$$u_2 = \frac{1}{\left\|\left(-\frac{1}{2}, \frac{1}{2}, 0, 1\right)\right\|}\left(-\frac{1}{2}, \frac{1}{2}, 0, 1\right) = \sqrt{\frac{2}{3}}\left(-\frac{1}{2}, \frac{1}{2}, 0, 1\right)$$

Entonces

$$\beta = \left\{ \frac{1}{\sqrt{2}}(1, 1, 0, 0),\quad \sqrt{\frac{2}{3}}\left(-\frac{1}{2}, \frac{1}{2}, 0, 1\right) \right\}$$

es una base ortonormal de W.

El vector $v_1 \in W$ será, según el análisis anterior

$$v_1 = (v \mid u_1)u_1 + (v \mid u_2)u_2 = \left(\frac{5}{\sqrt{2}}\right)\frac{1}{\sqrt{2}}(1, 1, 0, 0) + \sqrt{\frac{2}{3}}\left(\frac{9}{2}\right)\sqrt{\frac{2}{3}}\left(-\frac{1}{2}, \frac{1}{2}, 0, 1\right)$$

$$= \frac{5}{2}(1, 1, 0, 0) + 3\left(-\frac{1}{2}, \frac{1}{2}, 0, 1\right) = (1, 4, 0, 3)$$

El vector $v_2 \in W^\perp$ se obtiene por diferencia de v con v_1

$$v_2 = v - v_1 = (2, 3, -1, 4) - (1, 4, 0, 3) = (1, -1, -1, 1)$$

Entonces, se tiene

$$(2, 3, -1, 4) = (1, 4, 0, 3) + (1, -1, -1, 1)$$

en donde $(1, 4, 0, 3) \in W$ y $(1, -1, -1, 1) \in W^\perp$.

Por último se verá una aplicación del teorema 4.5 a la teoría de sistemas de ecuaciones lineales.

Considérese el sistema homogéneo de m ecuaciones lineales con n incógnitas

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0 \\ &\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0 \end{aligned} \qquad (S)$$

Sean los vectores c_i, definidos como

$$c_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \in \mathbf{R}^n \qquad i = 1, 2, \ldots, m$$

Con el producto interno canónico de $\mathbf{R}^n$, el sistema de ecuaciones puede verse como

$$\begin{aligned} (c_1 \mid X) &= 0 \\ (c_2 \mid X) &= 0 \\ &\vdots \\ (c_m \mid X) &= 0 \end{aligned}$$

en donde $X = (x_1, x_2, \ldots, x_n)$.

Sea W el subespacio de $\mathbf{R}^n$ generado por los vectores $c_1, c_2, \ldots, c_m$. Obsérvese que

$$X = (x_1, x_2, \ldots, x_n) \text{ es solución del sistema } (S) \Leftrightarrow X \in W^\perp$$

Es decir, que $W^\perp$ es precisamente el espacio solución del sistema (S).

Además es claro que los vectores c_i, $i = 1, 2, \ldots, m$, son los vectores línea de la matriz A del sistema (S), de modo que W es el espacio línea de A.

Como, según el teorema 4.5 se tiene

$$\mathbf{R}^n = W \oplus W^\perp$$

se debe tener entonces, por el corolario del teorema 5.6 del capítulo 3, que

$$n = \dim \mathbf{R}^n = \dim W + \dim W^\perp$$

de donde

$$\dim W^\perp = n - \dim W$$

Es decir, la dimensión del espacio solución del sistema (S) es n menos la dimensión del espacio línea de A (= rango de A). En el teorema 7.6 del capítulo 4 se había ya establecido este resultado con un argumento completamente distinto al aquí presentado.

APÉNDICE. EL MÉTODO DE MÍNIMOS CUADRADOS

En este apéndice se explotará la idea de "mejor aproximación" que se establecerá más adelante con el objeto de obtener el útil método de mínimos cuadrados, usado para ajustar curvas a una serie de datos experimentales.

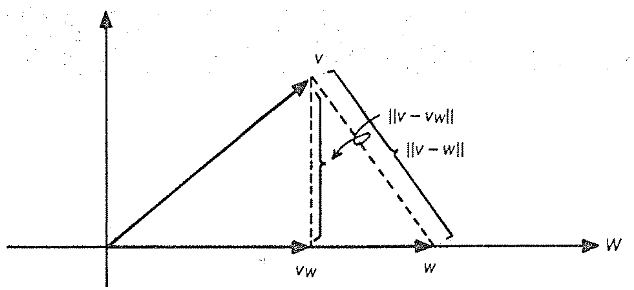
Sea V un espacio vectorial de dimensión finita con el producto interno $(\cdot \mid \cdot)$ y sea W un subespacio de él. Tómese un vector (arbitrario) fijo $v \in V$. Se dice que el vector $v_W \in W$ es *la mejor aproximación del vector v por vectores de W* si

$$\|v - v_W\| \leq \|v - w\| \qquad \forall\, w \in W$$

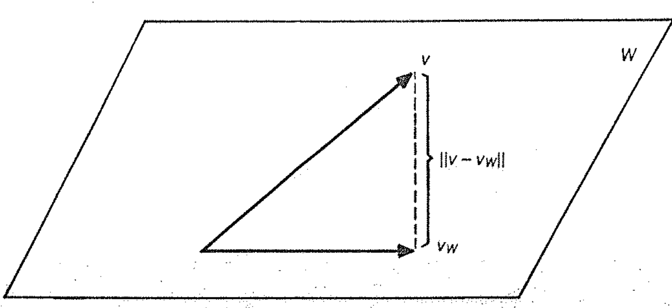
Al recordar que la norma de la diferencia entre dos vectores es la distancia entre ellos, la definición anterior establece que el vector $v_W \in W$ es la mejor aproximación del vector $v \in V$ por vectores de W si él es el vector de *W más próximo a v*. Es decir, de entre todos los vectores del subespacio W, el vector $v_W \in W$ es el vector cuya distancia a v es la menor posible.

Véanse un par de casos concretos, en los que la geometría del problema ayuda a descubrir al vector v_W.

Sea $V = \mathbf{R}^2$ con el producto interno canónico, y sea W el subespacio de V representado geométricamente por el eje x. Para un vector arbitrario $v \in \mathbf{R}^2$ es fácil ver (de la figura siguiente) que el vector v_W —la mejor aproximación de v por vectores de W— es la proyección ortogonal del vector v sobre el eje x, pues es claro que la distancia más corta de v a W es $\|v - v_W\|$.



Una situación similar se presenta en el caso del espacio vectorial $\mathbf{R}^3$ con el producto interno canónico, siendo W un plano que pasa por el origen: la proyección ortogonal del vector $v \in \mathbf{R}^3$ sobre W da la mejor aproximación del vector v por vectores de W.



Regrésese ahora al caso general. Se usará el teorema 4.5 para exhibir el vector v_W que mejor aproxima a un vector $v \in V$ por vectores de W. En ese teorema se vio que siendo V un espacio vectorial de dimensión finita con producto interno y W un subespacio de V, entonces V se podría presentar como la suma directa de W y su complemento ortogonal. Es decir, se podría escribir

$$V = W \oplus W^{\perp}$$

de modo que todo vector $v \in V$ podría ser escrito como

$$v = w_1 + w_2 \qquad \text{(A1)}$$

en donde $w_1 \in W$ y $w_2 \in W^{\perp}$. Al vector $w_1 \in W$ se le llama *proyección ortogonal* del vector v sobre el espacio W.

Se afirma que el vector w_1 es precisamente el vector v_W del análisis anterior. Es decir, que w_1 es la mejor aproximación de v por vectores de W.

En efecto, se probará a continuación que

$$\|v - w_1\| \leq \|v - w\| \qquad \forall\, w \in W \qquad \text{(A2)}$$

Según la expresión (A1) se tiene que para cualquier vector $w \in W$

$$\|v - w\|^2 = (w_1 + w_2 - w \mid w_1 + w_2 - w) = \|w_2\|^2 + \|w_1\|^2 + \|w\|^2 - 2(w_1 \mid w)$$

Pero

$$(w_1 \mid w) \leq |(w_1 \mid w)| \leq \|w_1\|\ \|w\|$$

y entonces

$$\|v - w\|^2 \geq \|w_2\|^2 + \|w_1\|^2 + \|w\|^2 - 2\|w_1\|\ \|w\| = \|w_2\|^2 + (\|w_1\| - \|w\|)^2$$
$$\geq \|w_2\|^2 = \|v - w_1\|^2$$

Es decir,

$$\|v - w_1\| \leq \|v - w\| \qquad \forall\, w \in W$$

lo que prueba la afirmación.

Se tiene entonces demostrado el siguiente teorema:

TEOREMA A1

Sea V un espacio vectorial de dimensión finita con producto interno $(\cdot \mid \cdot)$ y sea W un subespacio de V. Dado un vector $v \in V$, el vector $v_W \in W$ dado por

$$v_W = \sum_{i=1}^{k} (v \mid u_i) u_i \qquad \text{(A3)}$$

en donde $\beta = \{u_1, u_2, \ldots, u_k\}$ es una base ortonormal de W, da la mejor aproximación de v por vectores de W en el sentido de que

$$\|v - v_W\| \leq \|v - w\| \qquad \forall\, w \in W$$

La expresión del vector v_W en el teorema anterior se obtiene del análisis que sigue al teorema 4.5.

Considérese ahora un sistema no homogéneo de n ecuaciones lineales con n incógnitas

$$AX = B \qquad \text{(A4)}$$

en donde

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & \cdot\ \cdot\ \cdot & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Se sabe que este sistema es consistente si, y sólo si el vector $B \in \mathbf{R}^{m}$* pertenece al espacio columna de A (véase la demostración del teorema 7.7 capítulo 4), es de-

*Se considera al espacio $\mathbf{R}^m$ con el producto interno canónico.

cir si, y sólo si existen escalares $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ (la solución del sistema) tales que

$$B = \tilde{x}_1 v_1 + \tilde{x}_2 v_2 + \ldots + \tilde{x}_n v_n$$

en donde $v_j = (a_{1j}, a_{2j}, \ldots, a_{mj}) \in \mathbf{R}^m, j = 1, 2, \ldots, n$ son los vectores columna de A.

Llámese W al espacio columna de A. W es entonces un subespacio de $\mathbf{R}^m$.

Supóngase que el vector $B \in \mathbf{R}^m$ *no* pertenece al espacio columna de A. En tal caso, entonces el sistema (A4) no tiene solución. Es decir, no existe vector $X \in \mathbf{R}^n$ tal que $AX = B$ sea una identidad. En muchos casos de interés práctico es muy importante obtener "algo" que sea lo más parecido posible a una solución de este sistema. Es decir, interesa obtener un vector $X \in \mathbf{R}^n$ que sea una *solución aproximada* del sistema (inconsistente) $AX = B$.

Se explotarán las ideas analizadas al inicio de este apéndice para obtener este vector $\tilde{X}$ que aproxima a una solución del sistema (A4).

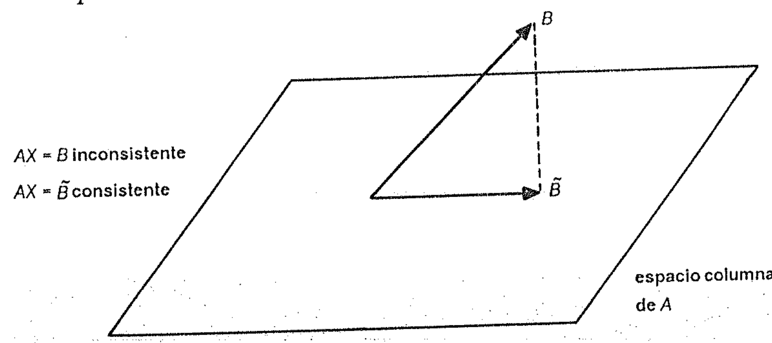
Como se ha visto, para poder obtener una solución del sistema $AX = B$ es necesario y suficiente que el vector $B \in \mathbf{R}^m$ pertenezca al espacio columna de A (al subespacio W de $\mathbf{R}^m$). La idea para obtener una solución aproximada del sistema inconsistente $AX = B$ es sustituir al vector $B \in \mathbf{R}^m$ que no pertenece a W, por el vector $\tilde{B} \in W$ que mejor lo aproxima.

Se trata entonces de considerar el sistema (consistente)

$$AX = \tilde{B} \tag{A5}$$

en donde $\tilde{B} \in W$ es la mejor aproximación del vector $B \in \mathbf{R}^m$ por vectores del espacio columna de A; así entonces el vector $\tilde{X} \in \mathbf{R}^n$ solución de este nuevo sistema, será una solución aproximada del sistema original.

Esquemáticamente, esta situación se ve como



Véanse funcionar estas ideas en algunos casos concretos.

Supóngase que al medir varias veces una cierta magnitud física x se obtienen los valores $b_1, b_2, \ldots, b_m$. Si $b_1 = b_2 = \ldots = b_m$ no hay problema alguno en afirmar que *el* valor de la magnitud x es el valor común de las m mediciones que se hicieron de ella. Sin embargo, esta situación en muy raras ocasiones ocurre. La situación más común es que los m valores $b_1, b_2, \ldots, b_m$ de x sean distintos. No obstante,

interesa reportar un valor de la magnitud x. ¿Cuál es este valor que se asigna a x? Contémplese este problema con la perspectiva del análisis que se ha hecho sobre sistemas inconsistentes y sus soluciones aproximadas.

Se tiene entonces el sistema inconsistente de m ecuaciones lineales con una incógnita x.

$$\begin{aligned} x &= b_1 \\ x &= b_2 \\ &\vdots \\ x &= b_m \end{aligned}$$

La matriz $m \times 1$ de este sistema es

$$A = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

El espacio columna de A es en este caso el subespacio W de dimensión 1 de $\mathbf{R}^m$ generado por el único vector columna de A

$$v = (1, 1, \ldots, 1)$$

Si el vector $B = (b_1, b_2, \ldots, b_m) \in \mathbf{R}^m$ perteneciera al espacio columna de A, el sistema $AX = B$ sería consistente. Obsérvese que éste es el caso en el que $b_1 = b_2 = \ldots = b_m$. Está uno suponiendo, sin embargo, que esto no ocurre. Sustitúyase entonces el vector $B \in \mathbf{R}^m$ por su mejor aproximación $\tilde{B} \in W$ por vectores de W = espacio columna de A = espacio generado por el vector $(1, 1, \ldots, 1)$.

Para poder usar la fórmula (A3) del teorema (A1) que da el vector $\tilde{B}$, se necesita una base ortonormal del espacio W. Es claro que en este caso sólo es necesario normalizar el vector v, de modo que el vector $u = v/\|v\|$ es una base ortonormal de W. Este vector es:

$$u = \frac{1}{\sqrt{m}}(1, 1, \ldots, 1)$$

El vector $\tilde{B}$ será entonces

$$\begin{aligned} \tilde{B} = (B \mid u)u &= \frac{1}{\sqrt{m}}(b_1 + b_2 + \ldots + b_m)\frac{1}{\sqrt{m}}(1, 1, \ldots, 1) \\ &= \frac{b_1 + b_2 + \ldots + b_m}{m}(1, 1, \ldots, 1) \\ &= (\tilde{b}_1, \tilde{b}_2, \ldots, \tilde{b}_m) \end{aligned}$$

en donde

$$\tilde{b}_i = \frac{b_1 + b_2 + \ldots + b_m}{m} \qquad i = 1, 2, \ldots, m$$

Se tiene entonces que el nuevo sistema consistente $AX = \tilde{B}$ tiene por solución

$$x_1 = x_2 = \ldots = x_m = \frac{b_1 + b_2 + \ldots + b_m}{m}$$

y por lo tanto, el valor que se reporta de la magnitud x de la que se obtuvo después de m mediciones los valores de $b_1, b_2, \ldots, b_m$ es:

$$x = \frac{b_1 + b_2 + \ldots + b_m}{m}$$

. . . la media aritmética de los valores $b_1, b_2, \ldots, b_m$.

Estúdiese ahora el caso en el que se tienen dos magnitudes x y y relacionadas entre sí. Se sabe teóricamente que la relación entre estas magnitudes es *lineal*, es decir, que existen constantes m y b tales que

$$y = mx + b$$

Después de efectuar algunas mediciones de estas magnitudes se obtienen los siguientes datos:

x	x_1	x_2	. . .	x_n
y	y_1	y_2	. . .	y_n

(A6)

Se trata entonces de calcular los valores de m y b para establecer la relación $y = mx + b$ entre x y y. Si las mediciones que se realizaron de x y y fueron efectuadas con precisión absoluta, sin error de ninguna naturaleza, bastarían las dos mediciones iniciales para concluir que

$$m = \frac{y_2 - y_1}{x_2 - x_1} \qquad \text{(A7)}$$

$$b = \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}$$

y entonces se podría comprobar que las mediciones restantes se ajustan a la recta $y = mx + b$, para estos valores de m y b.

Sin embargo, el caso que frecuentemente se encuentra es en el que aparecen errores en las mediciones de x y y, de modo que para calcular m y b se tiene que contestar ahora la pregunta: ¿cuál es la recta $y = mx + b$ que *mejor se ajusta* a los datos de la tabla (A6)?

Contémplese nuevamente este problema desde la perspectiva del análisis sobre sistemas inconsistentes y sus soluciones aproximadas.

Se tiene entonces un sistema de n ecuaciones con 2 incógnitas m y b.

$$\begin{aligned} x_1 m + b &= y_1 \\ x_2 m + b &= y_2 \\ &\vdots \\ x_n m + b &= y_n \end{aligned}$$

o bien,

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \cdot\ \cdot & \cdot \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

El caso en el que este sistema es consistente corresponde al caso en el que todas las mediciones de x y y de la tabla (A6) se ajustan a la recta $y = mx + b$ con m y b dados por las fórmulas (A7). El interés es analizar qué hacer, cuando este sistema *no* es consistente, para obtener los vectores de m y b que *mejor ajustan los datos* de la tabla (A6).

Se tiene entonces que el vector $Y = (y_1, y_2, \ldots, y_n) \in \mathbf{R}^n$ no se encuentra en el espacio columna de la matriz

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \cdot\ \cdot & \cdot \\ x_n & 1 \end{bmatrix}$$

Procédase entonces a obtener la mejor aproximación $\tilde{Y}$ de Y por vectores del espacio columna de A, para después resolver el sistema consistente $AX = \tilde{Y}$, en donde $X = (m, b)$.

En este caso, entonces el espacio columna de A es el subespacio W de $\mathbf{R}^n$ de dimensión 2 dado por

$$W = \mathcal{L}((1, 1, \ldots, 1), (x_1, x_2, \ldots, x_n))$$

Obténgase primeramente una base ortonormal $\beta = \{u_1, u_2\}$ de W, aplicando el proceso de Gram-Schmidt a los vectores $v_1 = (1, 1, \ldots, 1)$ y $v_2 = (x_1, x_2, \ldots, x_n)$ de la base de W.

Se tiene entonces que u_1 es

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{n}}(1, 1, \ldots, 1)_1$$

Para obtener el vector u_2 se calcula

$$v_2 - (v_2 \mid u_1)u_1 = (x_1, x_2, \ldots, x_n) - \frac{1}{\sqrt{n}}(x_1 + x_2 + \ldots + x_n)\frac{1}{\sqrt{n}}(1, 1, \ldots, 1)$$
$$= (x_1, x_2, \ldots, x_n) - \frac{x_1 + x_2 + \ldots + x_n}{n}(1, 1, \ldots, 1)$$

Llámese $\bar{x}$ a la media aritmética de los valores de x y $\bar{y}$ a la correspondiente de los valores de y.

Entonces

$$v_2 - (v_2 \mid u_1)u_1 = (x_1 - \bar{x}, x_2 - \bar{x}, \ldots, x_n - \bar{x})$$

la norma de este vector es:

$$\|v_2 - (v_2 \mid u_1)u_1\| = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

de modo que

$$u_2 = \frac{1}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}}(x_1 - \bar{x}, x_2 - \bar{x}, \ldots, x_n - \bar{x})$$

Procédase entonces a calcular $\tilde{Y}$ con la fórmula (A3)

$$\tilde{Y} = (Y \mid u_1)u_1 + (Y \mid u_2)u_2$$
$$= \frac{1}{\sqrt{n}}(y_1 + y_2 + \ldots + y_n)\frac{1}{\sqrt{n}}(1, 1, \ldots, 1)$$
$$+ \frac{1}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}}((x_1 - \bar{x})y_1 + \ldots + (x_n - \bar{x})y_n)$$
$$\cdot \frac{1}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}}(x_1 - \bar{x}, \ldots, x_n - \bar{x})$$
$$= \bar{y}(1, \ldots, 1) + \frac{\sum_{i=1}^{n} (x_i - \bar{x})y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2}(x_1 - \bar{x}, \ldots, x_n - \bar{x}) = (\tilde{y}_1, \ldots, \tilde{y}_n)$$

en donde

$$\tilde{y}_i = \bar{y} + \frac{\sum_{i=1}^{n} (x_i - \bar{x})y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2}(x_i - \bar{x}) \quad i = 1, 2, \ldots, n$$

El sistema $AX = \tilde{Y}$ es ahora consistente. La solución de éste es

$$m = \frac{\tilde{y}_2 - \tilde{y}_1}{x_2 - x_1} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$
$$b = \frac{x_1\tilde{y}_2 - x_2\tilde{y}_1}{x_1 - x_2} = \bar{y} - \frac{\sum_{i=1}^{n} (x_i - \bar{x})y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2} = \bar{y} - m\bar{x} \qquad \text{(A8)}$$

de modo entonces que la recta $y = mx + b$, con estos valores de m y b es la recta que mejor ajusta los datos de la tabla (A5). El procedimiento descrito anteriormente, por medio del cual se ajustan los datos de dos magnitudes relacionadas linealmente, según las fórmulas (A8), es conocido como *método de mínimos cuadrados*.

Véase, por último, un ejemplo numérico.

EJEMPLO 7 Al medir experimentalmente las magnitudes x y y, de las que se sabe que guardan entre sí una relación lineal, se obtuvieron los siguientes datos:

x	0.4	0.8	1.5	2.0	2.3	2.8
y	1.4	1.9	2.1	2.5	3.0	3.3

Se trata de obtener la recta que mejor ajuste estos datos.

Se tiene que

$$\bar{x} = \frac{0.4 + 0.8 + 1.5 + 2.0 + 2.3 + 2.8}{6} = 1.6333$$

$$\bar{y} = \frac{1.4 + 1.9 + 2.1 + 2.5 + 3.0 + 3.3}{6} = 2.3667$$

Al sustituir los correspondientes valores en las fórmulas (A8) se obtiene

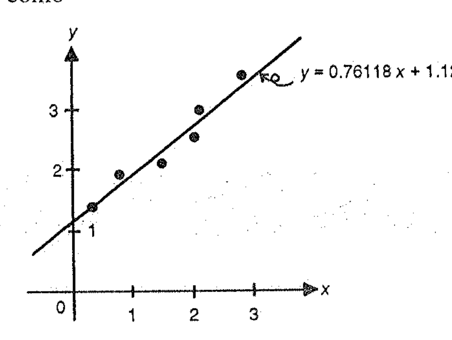
$$m = \frac{\sum_{i=1}^{6}(x_i - \bar{x})y_i}{\sum_{i=1}^{6}(x_i - \bar{x})^2} = \frac{3.17667}{4.17333} = 0.76118$$

$$b = \bar{y} - m\bar{x} = 1.1234$$

de modo que la recta que mejor ajusta los datos dados es:

$$y = 0.76118x + 1.1234$$

Gráficamente esto se ve como



EJERCICIOS. (SECCIÓN 4, CAPÍTULO 5)

1. Demuestre que

$$\beta = \left\{ \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \right\}$$

es una base ortonormal de $M_{2\times 2}$ con el producto interno canónico.

a) Escriba el vector

$$v = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$$

como combinación lineal de los vectores de β.

b) Escriba el vector

$$w = \begin{bmatrix} -1 & 3 \\ 2 & 1 \end{bmatrix}$$

como combinación lineal de los vectores de β.

c) Use las expresiones de v y w obtenidas en los incisos anteriores para calcular $(v \mid w)$ [véase teorema 4.1 inciso (3)].

d) Use las expresiones de v y w obtenidas en los incisos anteriores para calcular $\|v\|$ y $\|w\|$ [véase teorema 4.1 inciso (4)].

2. Considere el espacio vectorial $\mathbf{R}^2$ con el producto interno canónico. A partir de la base $\beta = \{(2, 2), (4, -2)\}$ de $\mathbf{R}^2$, construya una base ortonormal.

3. Considere el espacio vectorial $\mathbf{R}^3$ con el producto interno

$$(x \mid y) = x_1y_1 + 2x_2y_2 + x_3y_3$$

A partir de la base $\beta = \{(1, 1, 1), (2, 3, -1), (5, 2, -2)\}$ de $\mathbf{R}^3$ construya una base ortonormal.

4. En el espacio vectorial P_3 con el producto interno canónico considere la base

$$\beta = \{1, 1 + x, 1 + x + x^2, 1 + x + x^2 + x^3\}$$

A partir de esta base construya una base ortonormal de P_3.

5. Obtenga una base ortonormal para cada uno de los siguientes subespacios de $\mathbf{R}^3$ con el producto interno canónico

a) $\{(x, y, z) \mid 2x + 3y - z = 0\}$
b) $\{(x, y, z) \mid x = 3t, y = t, z = -2t, t \in \mathbf{R}\}$
c) $\{(x, y, z) \mid z = 0\}$
d) $\{(x, y, z) \mid x = y = z = t, t \in \mathbf{R}\}$

(*Sugerencia*: Obtenga primeramente una base cualquiera del subespacio correspondiente. Aplique luego el proceso de Gram-Schmidt.)

6. Obtenga una base ortonormal para cada uno de los siguientes subespacios de $\mathbf{R}^4$ (con el producto interno canónico).

a) $\{(x_1, x_2, x_3, x_4) \mid x_2 = x_3 = x_4 = 0\}$
b) $\{(x_1, x_2, x_3, x_4) \mid x_1 = x_3, x_2 = x_4\}$
c) $\{(x_1, x_2, x_3, x_4) \mid x_1 + x_2 + x_3 + x_4 = 0\}$
d) $\mathcal{L}((1, 1, 1, 1))$
e) $\mathcal{L}((1, 2, 3, 1), (2, 1, 2, 2))$
f) $\mathcal{L}((2, 3, -1, 2), (4, 6, -2, 4))$
g) $\mathcal{L}((1, 0, 1, 0), (0, 1, 0, 1), (-1, 0, 1, 0))$
h) $\mathcal{L}((1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1))$

7. Obtenga una base ortonormal del espacio solución de cada uno de los siguientes sistemas homogéneos de ecuaciones lineales (considere el producto interno canónico de $\mathbf{R}^n$).

a)
$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 - x_2 - x_3 &= 0 \\ 3x_1 + x_2 + x_3 &= 0 \end{aligned}$$

b)
$$\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 2x_1 - 4x_2 + 3x_3 &= 0 \\ -5x_1 + 13x_2 - 10x_3 &= 0 \end{aligned}$$

c)
$$\begin{aligned} 6x_1 + 4x_2 + 3x_3 + 5x_4 + 7x_5 &= 0 \\ 9x_1 + 6x_2 + 5x_3 + 7x_4 + 9x_5 &= 0 \\ 3x_1 + 2x_2 + x_3 + 3x_4 + 5x_5 &= 0 \\ 6x_1 + 4x_2 + 3x_4 + x_5 &= 0 \end{aligned}$$

8. Considere el espacio vectorial P_2 con el producto interno canónico. Obtenga una base ortonormal de P_2 que incluya a los vectores:

$$p(x) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}x, \quad q(x) = x^2$$

9. Determine el espacio vectorial $\mathbf{R}^3$ con el producto interno canónico. Obtenga una base ortonormal de $\mathbf{R}^3$ que incluya al vector $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$.

10. Calcule el espacio vectorial $M_{2\times 2}$ con el producto interno canónico. Obtenga una base ortonormal de $M_{2\times 2}$ que incluya a los vectores

$$v_1 = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \end{bmatrix} \qquad v_2 = \begin{bmatrix} 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 \end{bmatrix}$$

11. Determine el espacio vectorial P_3 con el producto interno

$$(p \mid q) = \int_{-1}^{1} p(x)q(x)dx$$

Aplique el proceso de Gram-Schmidt a la base $\beta = \{1, x, x^2, x^3\}$ para obtener una base ortonormal de P_4. Exprese el polinomio $1 + 2x - 3x^2 + x^3$ como combinación lineal de los elementos de esta base ortonormal.

① 12. Sea V un espacio vectorial de dimensión finita con el producto interno $(\cdot \mid \cdot)$ y sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base ortonormal de V. Considere el operador lineal T: $V \to V$. Demuestre que la matriz de T respecto de la base β es

$$[T]_\beta = ((T(v_j) \mid v_i))_{i,j=1,\ldots,n}$$

13. Compruebe que si $A \in M_{n\times n}$ es una matriz ortogonal, su transpuesta A^t también es ortogonal.

14. Suponga que las matrices cuadradas A y B de orden n son ortogonales. Demuestre que AB y BA son también ortogonales.

15. Suponga que la matriz $A \in M_{n\times n}$ es ortogonal y además es simétrica. Demuestre que $A^2 = I$.

② 16. Sea A una matriz cuadrada de orden n, pruebe que las siguientes afirmaciones acerca de A son equivalentes.

a) A es una matriz ortogonal.

b) Los vectores línea de A forman una base ortonormal de $\mathbf{R}^n$.

c) Los vectores columna de A forman una base ortonormal de $\mathbf{R}^n$.

① 17. Demuestre que si la matriz ortogonal $A \in M_{n\times n}$ es triangular, entonces es diagonal.

18. Calcule α y β de modo que la matriz

$$A = \begin{bmatrix} \alpha & 1 \\ 1 & \beta \end{bmatrix}$$

sea ortogonal.

19. Compruebe que para cualquier valor de θ las matrices

$$A = \begin{bmatrix} \operatorname{sen}\theta & \cos\theta \\ \cos\theta & -\operatorname{sen}\theta \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \operatorname{sen}\theta & \cos\theta \\ 0 & \cos\theta & -\operatorname{sen}\theta \end{bmatrix}$$

son ortogonales.

20. Con base en el ejercicio 1, argumente por qué la matriz

$$A = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{bmatrix}$$

es una matriz ortogonal.

(*Sugerencia*: use el teorema 4.3.)

21. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$. Sea W un subespacio de dimensión finita de V y $\beta = \{v_1, v_2, \ldots, v_k\}$ una base de W, demuestre que las siguientes afirmaciones acerca del vector $u \in V$ son equivalentes.

a) $u \in W^\perp$ b) $(u \mid v_i) = 0, \quad i = 1, 2, \ldots, k.$

22. Use el ejercicio anterior para demostrar que si $v = (a, b, c)$ es un vector no nulo en $\mathbf{R}^3$ (con el producto interno canónico), entonces el complemento ortogonal del subespacio W generado por v es:

$$W^\perp = \{(x, y, z) \mid ax + by + cz = 0\}$$

Interprete geométricamente.

23. En el espacio vectorial $\mathbf{R}^3$ con el producto interno canónico considere los vectores $v_1 = (1, 1, 1)$ y $v_2 = (2, 1, -1)$. Sea $W = \mathcal{L}(v_1, v_2)$. Demuestre que el complemento ortogonal de W es una recta que pasa por el origen. Determine la ecuación de esta recta. Interprete geométricamente.

24. ¿Cierto o falso? Si v_1 y v_2 son dos vectores en $\mathbf{R}^3$ con el producto interno canónico, el complemento ortogonal del subespacio de $\mathbf{R}^3$ generado por v_1 y v_2 es una recta que pasa por el origen.

25. Sean a, b y c tres números reales no nulos, determine el complemento ortogonal del subespacio W de $\mathbf{R}^3$ (con el producto interno canónico) dado por

$$W = \{(x, y, z) \mid ax + by + cz = 0\}$$

Interprete geométricamente.

26. Considere el espacio $M_{2\times 2}$ con el producto interno canónico. Determine el complemento ortogonal de los siguientes subespacios de $M_{2\times 2}$.

a) $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \,\middle|\, b = c = 0\right\}$

b) $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \,\middle|\, a + b + c + d = 0\right\}$

c) $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \,\middle|\, a = 0\right\}$

d) $\{A \mid A = A^t\}$

e) $\{A \mid A = -A^t\}$

① 27. Considere el espacio P_3 con el producto interno $(p \mid q) = \int_{-1}^{1} p(x)q(x)dx$. Determine el complemento ortogonal del subespacio de P_3 formado por los polinomios constantes.

28. Encuentre el complemento ortogonal del espacio solución del sistema homogéneo de ecuaciones lineales

$$2x_1 - x_2 + 5x_3 + 7x_4 = 0$$
$$6x_1 - 3x_2 + 8x_3 = 0$$
$$4x_1 - 2x_2 + 6x_3 + 2x_4 = 0$$

(considere el producto interno canónico).

② 29. Sea V un espacio vectorial de dimensión finita con el producto interno $(\cdot \mid \cdot)$. Sea W un subespacio de V y sea U otro subespacio de V tal que $V = W \oplus U$. ¿Se puede concluir que $U = W^{\perp}$?

② 30. Sea V un espacio vectorial de dimensión finita con el producto interno $(\cdot \mid \cdot)$. Sea W un subespacio de V. Escriba $V = W \oplus W^{\perp}$. Considere la transformación $T: V \to V$ definida como: para $v \in V$, $v = w_1 + w_2$, $w_1 \in W$, $W_2 \in W^{\perp}$, $T(v) = w_1$.

a) Demuestre que T es una transformación lineal.
b) Compruebe que $(T(v_1) \mid v_2) = (v_1 \mid T(v_2))\ \forall\, v_1, v_2 \in V$.
c) Describa el operador T^2.
d) Interprete geométricamente la acción del operador T en el caso de $\mathbf{R}^2$ y $\mathbf{R}^3$ con el producto interno canónico.

Nota: Al operador T se le llama *proyección* de V sobre el subespacio W.

31. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$, demuestre que

a) $\{0\}^{\perp} = V$
b) $V^{\perp} = \{0\}$

32. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$. Sean W_1 y W_2 dos subespacios de V tales que $W_1 \subset W_2$. Demuestre que $(W_2^{\perp} \subset W_1^{\perp})$.

33. Sea V un espacio vectorial de dimensión finita con el producto interno $(\cdot \mid \cdot)$. Sea W un subespacio de V. Demuestre que $(W^{\perp})^{\perp} = W$.

34. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$. Sean W_1 y W_2 dos subespacios de V, pruebe que

a) $(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$
b) $(W_1 \cap W_2)^{\perp} = W_1^{\perp} + W_2^{\perp}$

35. Sea V un espacio vectorial de dimensión finita con el producto interno $(\cdot \mid \cdot)$. Sean W_1 y W_2 dos subespacios de V tales que $W_1^{\perp} = W_2^{\perp}$, demuestre que $W_1 = W_2$.

36. Para cada uno de los siguientes subespacios W de $\mathbf{R}^3$, escriba el vector v dado como una suma de un vector de W y otro de $W^{\perp}$ (considere el producto interno canónico de $\mathbf{R}^3$).

a) $W = \{(x, y, z) \mid x + 2y - z = 0\}$ $v = (1, 3, 1)$
b) $W = \{(x, y, z) \mid x = y = 0\}$ $v = (2, 1, 0)$
c) $W = \{(x, y, z) \mid x = y = z = t, t \in \mathbf{R}\}$ $v = (3, 1, 2)$
d) $W = \{(x, y, z) \mid x + y + z = 0\}$ $v = (1, 1, -2)$

37. Para cada uno de los subespacios W de $\mathbf{R}^4$, escriba el vector v dado como una suma de un vector de W y otro de $W^{\perp}$ (considere el producto interno canónico de $\mathbf{R}^4$).

a) $W = \{(x_1, x_2, x_3, x_4) \mid x_1 = x_2\}$ $v = (1, 3, 1, 1)$
b) $W = \{(x_1, x_2, x_3, x_4) \mid x_1 + x_2 + x_3 + x_4 = 0\}$ $v = (2, 1, 1, 0)$
c) $W = \mathscr{L}((1, 3, 1, 1), (2, -1, 0, 1))$ $v = (1, 3, 1, 0)$
d) $W = \mathscr{L}((2, 1, -1, 0), (3, 1, 0, 1))$ $v = (2, 1, 0, 1)$

38. Sea W el espacio solución (subespacio de $\mathbf{R}^4$) del sistema homogéneo de ecuaciones lineales

$$2x_1 + x_2 + x_3 + 3x_4 = 0$$
$$4x_1 + 4x_2 + 4x_3 - 8x_4 = 0$$
$$3x_1 + 3x_2 + 3x_3 - 6x_4 = 0$$

Escriba el vector $(7, -4, -1, 2)$ como una suma de un vector en W y otro en $W^{\perp}$.

EJERCICIOS (APÉNDICE. SECCIÓN 4, CAPÍTULO 5)

39. Encuentre la recta $y = mx + b$ que mejor ajuste a los siguientes datos:

x	0	1.3	1.8	2.2
y	0.4	3.5	5	5.5

40. Bajo ciertas condiciones, se puede considerar que la reacción de oxidación de tiourea por hexaciano ferrato (III) en medio alcalino

$$NH_2CSNH_2 + 10(OH)^- + 8(Fe(CN)_6)^{3-} \to NH_2CONH_2 + SO_4^{2-} + 8(Fe(CN)_6)^{4-} + 5H_2O$$

es una reacción de primer orden, para la cual se cumple que

$$C = C_0 e^{-kt} \qquad \text{(I)}$$

en donde C es la concentración de hexaciano a un tiempo t, C_0 es la concentración inicial de este reactivo (concentración al tiempo $t = 0$) y k es la llamada constante de velocidad de reacción, parámetro cuyo conocimiento reporta una gran cantidad de información sobre la misma.

Obsérvese que la expresión (I) puede ser escrita como

$$\ln C = \ln C_0 - kt$$

quedando ahora clara la dependencia lineal existente entre $\ln C$ y t.

El valor de la constante de velocidad de reacción k es una función del pH de la solución reaccionante, de la fuerza iónica y de la temperatura a la que se efectúa la reacción. A continuación se muestran varias tablas de datos obtenidos experimentalmente a diferentes condiciones de reacción en las que se dan los valores de C (medido en moles/litro) a diferentes tiempos durante la reacción

pH = 9.0
Fza. iónica = normal
Temperatura = 40 °C

t (mín)	$C(\times 10^4)(M)$
3	7.847
15	7.406
40	5.721
80	3.486
100	2.867
120	2.704

pH = 9.0
Fza. iónica = normal
Temperatura = 50 °C

t (mín)	$C(\times 10^4)(M)$
3	7.406
15	6.055
40	2.956
80	1.800
100	1.586
120	1.485

pH = 11.0
Fza. iónica = normal
Temperatura = 40 °C

t (mín)	$C(\times 10^4)(M)$
2	6.217
7	4.802
9	4.032
12	3.336
16	2.485
19	1.791

pH = 11.0
Fza. iónica = normal
Temperatura = 50 °C

t (mín)	$C(\times 10^4)(M)$
1	5.911
2	4.550
3	3.033
4	2.152
5.5	1.676
8	1.356

Calcule la constante de velocidad de reacción en cada caso.

41. Se sabe que las magnitudes x y y están relacionadas cuadráticamente, es decir, que existen constantes a, b y c tales que $y = ax^2 + bx + c$.
Al realizar algunas mediciones experimentales de estas magnitudes se obtuvieron los siguientes datos:

x	0.2	0.7	1.2	2
y	1.6	2.2	6	16

Encuéntrese la parábola que mejor ajuste estos datos, formando primeramente el sistema (inconsistente) de ecuaciones lineales en las incógnitas a, b y c

$$ax_i^2 + bx_i + c = y_i \qquad i = 1, 2, \ldots, 6$$

sustituyendo luego el vector $Y = (y_1, y_2, y_3, y_4) \in \mathbb{R}^4$ por su mejor aproximación $\tilde{Y}$ por medio de vectores del espacio columna de la matriz A del sistema y resolviendo finalmente el sistema $AX = \tilde{Y}$, en donde $X = (a, b, c)$.

② 42. En este ejercicio se muestra un acercamiento distinto al presentado en este apéndice, al método de mínimos cuadrados.

a) Supóngase que al medir la magnitud x n veces se obtuvieron los valores $x_1, x_2, \ldots, x_n$. El valor $\tilde{x}$ que mejor representa a x es, según el método de mínimos cuadrados, aquel que hace que la suma de los cuadrados de las desviaciones (las diferencias) de $\tilde{x}$ respecto de cada uno de los valores $x_1, x_2, \ldots, x_n$, sea mínimo. Es decir, si S es la suma mencionada, entonces

$$S = \sum_{i=1}^{n} (x_i - \tilde{x})^2 = \begin{array}{l}\text{suma de los cuadrados de las}\\ \text{desviaciones de } \tilde{x} \text{ respecto de } x_i\\ \text{(mínima)}\end{array}$$

Demuéstrese que el valor de x que hace que S sea mínima es

$$\tilde{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

(*Sugerencia*: impóngase la condición de extremo local a S y resuelva para $\tilde{x}$.)

b) Supóngase ahora que las magnitudes x y y están relacionadas linealmente. Al medir estas magnitudes experimentalmente se obtuvieron los siguientes datos:

x	x_1	x_2	...	x_n
y	y_1	y_2	...	y_n

Se quieren ajustar estos datos a una recta $y = mx + b$. La recta que mejor ajusta estos datos es, según el método de mínimos cuadrados, aquella que tiene los valores de m y b tales que la suma de los cuadrados de las desviaciones de $\tilde{y}_i = mx_i + b$ respecto de y_i sea mínima; es decir, si m y b son tales que

$$S = \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2 = \sum_{i=1}^{n} (y_i - mx_i - b)^2$$

es mínima. Obsérvese que ahora S es una función de las dos variables m y b. Imponiendo las condiciones de extremo relativo

$$\frac{\partial S}{\partial m} = 0 \qquad \frac{\partial S}{\partial b} = 0$$

y resolviendo para m y b, demuestre que los valores de m y b que logran que S sea mínima son precisamente los que se obtuvieron en las fórmulas (A8).

5. TRANSFORMACIONES ORTOGONALES

Sea $T: V \to V$ una transformación lineal del espacio vectorial V de dimensión finita con producto interno, en sí mismo. Se dice que T es una *transformación ortogonal* si

$$\|T(v)\| = \|v\| \qquad \forall\, v \in V$$

Es decir, $T: V \to V$ es una *transformación ortogonal* si T *preserva la norma*, en el sentido que la norma de la imagen de cualquier vector $v \in V$ bajo T es igual a la norma de v.

Obsérvese que si T es ortogonal, entonces T es inversible, pues si $v \in V$ es tal que $T(v) = 0$ se tiene

$$\|v\| = \|T(v)\| = \|0\| = 0$$

de donde $v = 0$. Entonces $Ker\, T = \{0\}$, lo que equivale a decir que T es inversible.

El siguiente teorema da condiciones equivalentes al hecho de ser $T: V \to V$ una *transformación ortogonal.*

TEOREMA 5.1

Las siguientes afirmaciones son equivalentes.

1) La transformación $T: V \to V$ es ortogonal.
2) T preserva el producto interno. Es decir, para cualesquiera $u, v \in V$ se tiene

$$(T(u) \mid T(v)) = (u \mid v)$$

3) T transforma bases ortonormales en bases ortonormales. Es decir, si $\beta_1 = \{v_1, v_2, \ldots, v_n\}$ es una base ortonormal en V, entonces $\beta_2 = \{T(v_1), T(v_2), \ldots, T(v_n)\}$ es también una base ortonormal de V.
4) La matriz que representa a T respecto de una base ortonormal de V es una matriz ortogonal.

DEMOSTRACIÓN Se probará que 1) ⇒ 2) ⇒ 3) ⇒ 4) ⇒ 1).

1) ⇒ 2). Primeramente, obsérvese que para $u, v \in V$ se tiene

$$\|u + v\|^2 = (u + v \mid u + v) = (u \mid u) + (u \mid v) + (v \mid u) + (v \mid v)$$
$$= \|u\|^2 + 2(u \mid v) + \|v\|^2$$

de donde

$$(u \mid v) = \frac{1}{2}(\|u + v\|^2 - \|u\|^2 - \|v\|^2)$$

Entonces, siendo T ortogonal se tiene que

$$(T(u) \mid T(v)) = \frac{1}{2}(\|T(u) + T(v)\|^2 - \|T(u)\|^2 - \|T(v)\|^2) = \frac{1}{2}(\|T(u + v)\|^2 - \|T(u)\|^2 - \|T(v)\|^2)$$
$$= \frac{1}{2}(\|u + v\|^2 - \|u\|^2 - \|v\|^2) = (u \mid v)$$

y así entonces T preserva el producto interno.

2) ⇒ 3). Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base ortonormal de V, entonces

$$(T(v_i) \mid T(v_j)) = (v_i \mid v_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Es decir, el conjunto $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ es un conjunto ortonormal de V. Como $dim\, V = n$, este conjunto es de hecho una base ortonormal de V.

3) ⇒ 4). Sea $\beta_1 = \{v_1, v_2, \ldots, v_n\}$ una base ortonormal de V, y sea $A = [T]_{\beta_1} = (a_{ij})_{i,j=1,\ldots,n}$. Entonces como se sabe

$$T(v_j) = \sum_{i=1}^{n} a_{ij}v_i \qquad j = 1, 2, \ldots, n$$

Como $\beta_2 = \{T(v_1), T(v_2), \ldots, T(v_n)\}$ es también (por hipótesis) una base ortonormal de V, se puede ver (gracias a la expresión anterior para $T(v_j)$) a la matriz A como la matriz de cambio de base de β_2 a β_1, pues en su j-ésima columna aparece el vector de coordenadas de $T(v_j)$ en términos de la base β_1. Por tanto, según el teorema 4.3, A es una matriz ortogonal, como se quería.

4) ⇒ 1). Sea $\beta = \{v_1, v_2, \ldots, v_n\}$ una base ortonormal de V y sea $A = [T]_\beta = (a_{ij})_{i,j=1,\ldots,n}$. Se sabe (por hipótesis) que A es una matriz ortogonal, esto es, $A^tA = I$. Entonces

$$\begin{array}{l}\text{Elemento de la } i\text{-ésima línea y} \\ j\text{-ésima columna de } A^tA\end{array} = \sum_{k=1}^{n} a_{ki}a_{kj} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Sea ahora u un vector cualquiera de V, se puede escribir

$$u = c_1v_1 + c_2v_2 + \ldots + c_nv_n = \sum_{j=1}^{n} c_jv_j$$

Entonces

$$T(u) = T\left(\sum_{j=1}^{n} c_jv_j\right) = \sum_{j=1}^{n} c_jT(v_j)$$

Pero como $A = [T]_\beta$, se tiene

$$T(v_j) = \sum_{i=1}^{n} a_{ij}v_i$$

por lo que

$$T(u) = \sum_{j=1}^{n} c_jT(v_j) = \sum_{j=1}^{n} c_j\sum_{i=1}^{n} a_{ij}v_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}c_j\right)v_i$$

expresión que da al vector $T(u) \in V$ como una combinación lineal de los vectores de la base ortonormal β.

En vista del teorema 4.1 [apartados (2) y (4)] se puede concluir que

$$(T(u) \mid v_i) = \sum_{j=1}^{n} a_{ij}c_j$$

y por lo tanto,

$$\|Tu\|^2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ir}a_{ij}\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r=1}^{n} a_{ij}a_{ir}c_jc_r$$

$$= \sum_{j=1}^{n}\left[\sum_{r=1}^{n}\underbrace{\left(\sum_{i=1}^{n} a_{ir}a_{ij}\right)}_{=\begin{cases}1 & \text{si } r=j\\ 0 & \text{si } r\neq j\end{cases}} c_r\right]c_j = \sum_{j=1}^{n} c_j^2 = \|u\|^2$$

Entonces $\|Tu\| = \|u\|$, quedando así probado que T es ortogonal.

Q.E.D.

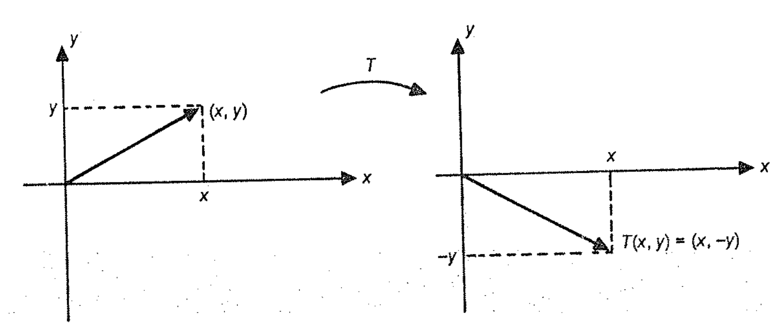
EJEMPLO 1 Considere, por ejemplo, en el espacio $\mathbf{R}^2$ con el producto interno canónico la transformación T: $\mathbf{R}^2 \to \mathbf{R}^2$ dado por

$$T(x, y) = (x, -y)$$

T es una transformación ortogonal pues

$$\|T(x, y)\| = \|(x, -y)\| = \sqrt{x^2 + (-y)^2} = \sqrt{x^2 + y^2} = \|(x, y)\| \qquad \forall (x, y) \in \mathbf{R}^2$$

Geométricamente, T es una reflexión respecto del eje x.



Obsérvese que si $v_1 = (x_1, y_1)$ y $v_2 = (x_2, y_2)$ se tiene

$$(Tv_1 \mid Tv_2) = ((x_1, -y_1) \mid (x_2 - y_2)) = (x_1)(x_2) + (-y_1)(-y_2) = x_1x_2 + y_1y_2 = (v_1 \mid v_2)$$

Es decir, que T preserva el producto interno, tal como lo asegura el teorema 5.1. También, según este teorema, la matriz de T respecto de la base canónica (que es una base ortonormal de $\mathbf{R}^2$) será una matriz ortogonal. Esta matriz es:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

En efecto,

$$A^tA = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Una transformación lineal T: $V \to V$ (V de dimensión finita) que preserva distancias, es decir tal que

$$d(T(v), T(u)) = d(v, u) \qquad \forall\, v, u \in V$$

se llama *transformación isométrica* (o isometría). El siguiente teorema dice que la propiedad de una transformación lineal, ser ortogonal y ser isométrica, son equivalentes.

TEOREMA 5.2 La transformación lineal T: $V \to V$, en donde V es un espacio vectorial de dimensión finita con producto interno, es una transformación ortogonal si, y sólo si es isométrica.

DEMOSTRACIÓN Supóngase que T es ortogonal. Entonces, para $v, u \in V$, $d(T(v), T(u)) = \|T(v) - T(u)\| = \|T(v - u)\| = \|v - u\| = d(v, u)$ lo que muestra que T es isométrica.

Recíprocamente, supóngase que T es isométrica. Entonces,

$$\|T(v)\| = \|T(v) - 0\| = \|T(v) - T(0)\| = d(T(v), T(0)) = d(v, 0) = \|v - 0\| = \|v\|$$

lo que muestra que T es ortogonal

Q.E.D.

En la sección 6 del capítulo anterior se había definido (definición 6.4) el determinante de una transformación lineal T: $V \to V$ (V de dimensión finita) como el determinante de la matriz que representa a T en cualquier base de V. Si se toma una base ortonormal β de V, el determinante de la *transformación ortogonal* T: $V \to V$ será el determinante de la matriz ortogonal $A = [T]_\beta$. Pero, según el teorema 4.4, $det\, A$ es 1 o -1. Entonces $det\, T$ es 1 o -1. Se ha probado así el siguiente teorema:

TEOREMA 5.3 Si T: $V \to V$ es una transformación ortogonal, entonces

$$det\, T = 1 \text{ o } -1$$

A una transformación ortogonal $T: \mathbf{R}^n \to \mathbf{R}^n$, tal que $det\ T = 1$, se llama *rotación*. Esta definición está inspirada en el caso particular $T: \mathbf{R}^2 \to \mathbf{R}^2$, T ortogonal, $det\ T = 1$, en el cual el efecto geométrico de T sobre un vector $v \in \mathbf{R}^2$ es una rotación de éste por un ángulo θ, como se verá a continuación.

Sea $T: \mathbf{R}^2 \to \mathbf{R}^2$ una transformación ortogonal tal que $det\ T = 1$. Sea $\beta = \{v_1, v_2\}$ una base ortonormal de $\mathbf{R}^2$, y se escribe $T(v_1) = (a, b)$, $T(v_2) = (c, d)$.

Entonces la matriz de T respecto de la base β es

$$[T]_\beta = A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

Según el teorema 5.1, A es una matriz ortogonal, es decir $A^t = A^{-1}$. Se sabe que

$$A^{-1} = \frac{1}{det\ A}\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$$

Como $det\ A = 1$, se tiene entonces

$$A^t = \begin{bmatrix} d & b \\ c & d \end{bmatrix} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix} = A^{-1}$$

de donde se obtiene que $a = d$ y $b = -c$. Es decir,

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

en donde a y b deben ser tales que

$$det\ A = a^2 + b^2 = 1$$

Al escribir el vector $(a, b) \in \mathbf{R}^2$ en coordenadas polares se tiene

$$a = r\cos\theta$$

$$b = r\,\text{sen}\,\theta$$

Pero,

$$r = \sqrt{a^2 + b^2} = 1$$

Entonces,

$$a = \cos\theta$$

$$b = \text{sen}\,\theta$$

de modo que la matriz A de T es:

$$A = \begin{bmatrix} \cos\theta & -\text{sen}\,\theta \\ \text{sen}\,\theta & \cos\theta \end{bmatrix}$$

y entonces la transformación $T: \mathbf{R}^2 \to \mathbf{R}^2$ puede describirse como

$$T(x, y) = A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\text{sen}\,\theta \\ \text{sen}\,\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta & -y\,\text{sen}\,\theta \\ x\,\text{sen}\,\theta & -\ y\cos\theta \end{bmatrix}$$

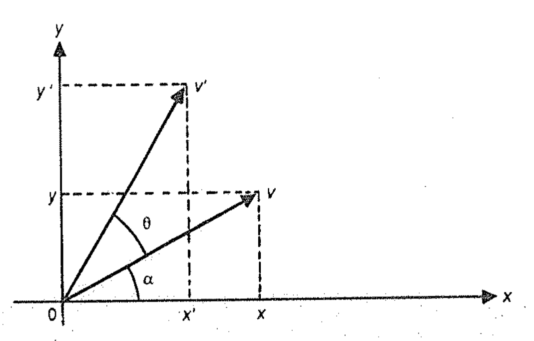
Por lo tanto, el vector $(x', y') = T(x, y)$ es tal que

$$x' = x\cos\theta - y\,\text{sen}\,\theta$$

$$y' = x\,\text{sen}\,\theta + y\cos\theta$$

Es fácil ver que (x', y') es un vector de $\mathbf{R}^2$ que se obtiene del vector $(x, y) \in \mathbf{R}^2$ girando éste un ángulo θ.

En efecto, considérese la siguiente figura:



Llámese K a la norma de v. Es decir, $K = ||v|| = ||v'||$.

De la figura se tiene

$$\text{sen}\,\alpha = \frac{y}{K} \qquad \cos\alpha = \frac{x}{K}$$

$$\text{sen}\,(\alpha + \theta) = \frac{y'}{K} \qquad \cos(\alpha + \theta) = \frac{x'}{K}$$

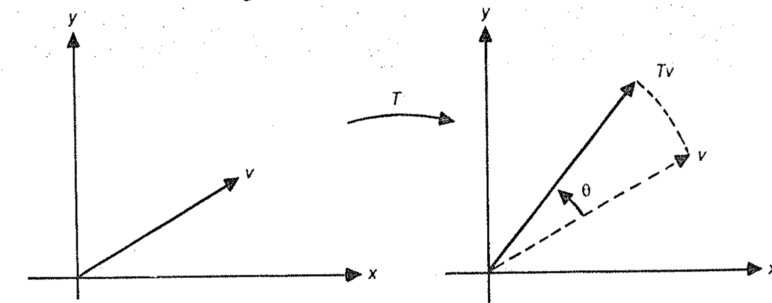
Entonces,

$$x' = K\cos(\alpha + \theta) = (K\cos\alpha)\cos\theta - (K\,\text{sen}\,\alpha)\,\text{sen}\,\theta = x\cos\theta - y\,\text{sen}\,\theta$$

$$y' = K\,\text{sen}\,(\alpha + \theta) = (K\,\text{sen}\,\alpha)\cos\theta + (K\cos\alpha)\,\text{sen}\,\theta = y\cos\theta + x\,\text{sen}\,\theta$$

que son las expresiones que definen a las coordenadas de $T(x, y)$.

Se tiene entonces la siguiente situación geométrica:



Esto justifica entonces que a la transformación ortogonal T: $\mathbf{R}^n \to \mathbf{R}^n$ tal que $det\ T = 1$, se le llama rotación (se puede demostrar —véanse ejercicios al final de esta sección— que también en el caso T: $\mathbf{R}^3 \to \mathbf{R}^3$, T ortogonal, $det\ T = 1$, el efecto geométrico de T sobre el vector $v \in \mathbf{R}^3$ —¡el cual puede dibujarse!— es también una rotación).

Véase ahora el caso de la transformación ortogonal T: $\mathbf{R}^2 \to \mathbf{R}^2$ tal que $det\ T = -1$.

Repítase el mismo argumento del caso $det\ T = 1$, hasta que se llegue a la fórmula de A^{-1}.

Como en este caso $det\ A = -1$, se tiene

$$A^t = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -d & c \\ b & -a \end{bmatrix} = A^{-1}$$

de donde $a = -d$, $c = b$. Es decir, que la matriz de T respecto de la base β es:

$$A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

en donde $det\ A = -a^2 - b^2 = -1$, o sea $a^2 + b^2 = 1$.

Al escribir el vector (a, b) en coordenadas polares, se ve que A puede escribirse como

$$A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}$$

Se había visto ya que la transformación S: $\mathbf{R}^2 \to \mathbf{R}^2$, $S(x, y) = (x, -y)$, o bien, $S(v) = Bv$, en donde

$$B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

es una transformación ortogonal la cual representa geométricamente una reflexión respecto del eje x (véase ejemplo después del teorema 5.1). Obsérvese que

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix} = A$$

La primera matriz que aparece en este producto es la matriz que corresponde a la transformación ortogonal R: $\mathbf{R}^2 \to \mathbf{R}^2$, en el caso $det\ R = 1$, esto es, es una matriz de rotación. Por lo tanto, si se considera la composición

$$\mathbf{R}^2 \xrightarrow[\text{(reflexión)}]{S} \mathbf{R}^2 \xrightarrow[\text{(rotación)}]{R} \mathbf{R}^2$$

$R \circ S$

se ve que la matriz que representa a $R \circ S$ (que es el producto de la matriz que representa a R por la matriz que representa a S) es precisamente A = matriz que representa a la transformación ortogonal T: $\mathbf{R}^2 \to \mathbf{R}^2$, en el caso de $det\ T = -1$.

Se puede concluir entonces que

$$T = R \circ S$$

Es decir, en el caso $det\ T = -1$, la transformación ortogonal T: $\mathbf{R}^2 \to \mathbf{R}^2$ es geométricamente una reflexión respecto del eje x seguida de una rotación.

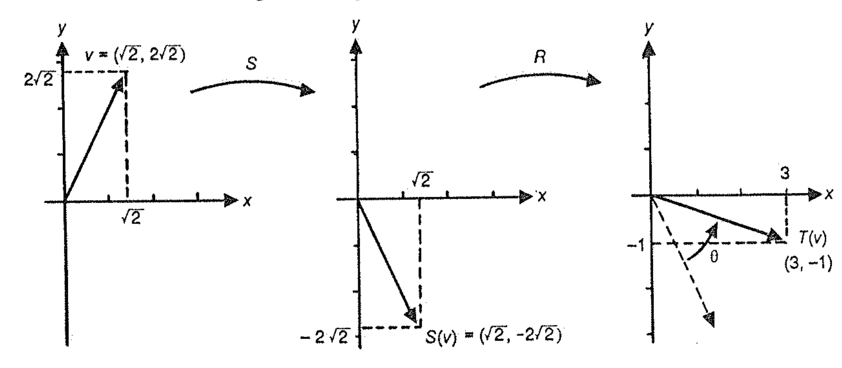
EJEMPLO 2 Por ejemplo, si T: $\mathbf{R}^2 \to \mathbf{R}^2$ es la transformación ortogonal $T(v) = Av$ en donde

$$A = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix}$$

y $v = (\sqrt{2}, 2\sqrt{2})$, se ve que

$$T(2, 2\sqrt{2}) = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 2\sqrt{2} \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

En este caso, $det\ T = -1$, de modo que el efecto geométrico de T sobre v es primero una reflexión de éste respecto del eje x y luego una rotación (en este caso de $\pi/4$), como se muestra en la siguiente figura:



EJERCICIOS (SECCIÓN 5, CAPÍTULO 5)

1. Considere la transformación lineal T: $\mathbf{R}^2 \to \mathbf{R}^2$ para la cual

$$T\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

$$T\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Demuestre que T es una transformación ortogonal.

2. Considere la transformación lineal $T: M_{2\times 2} \to M_{2\times 2}$ para la cual

$$T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \qquad T\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & -1/2 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{bmatrix} \qquad T\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}$$

Compruebe que T es una transformación ortogonal.
(*Sugerencia*: véase ejercicio 1 de la sección anterior.)

3. Sea $T: \mathbf{R}^3 \to \mathbf{R}^3$ la transformación lineal $T(X) = AX$, en donde

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$

Demuestre que T es una transformación ortogonal.
(*Sugerencia*: Véase ejercicio 18 de la sección anterior).

4. Dé tres argumentos distintos que muestren que la transformación identidad $Id: V \to V$, en donde V es un espacio vectorial de dimensión finita con el producto interno $(\cdot \mid \cdot)$, es una transformación ortogonal.
(*Sugerencia*: véase teorema 5.1.)

5. Sea V un espacio vectorial de dimensión 3 con producto interno $(\cdot \mid \cdot)$.
Sea $\beta = \{v_1, v_2, v_3\}$ una base ortonormal de V. Suponga que

$$(T(v_i) \mid T(v_j)) = (v_i \mid v_j) \qquad i, j = 1, 2, 3$$

Pruebe que T es una transformación ortogonal.

6. Determine el vector $(a, b) \in \mathbf{R}^2$ tal que la transformación $T: \mathbf{R}^2 \to \mathbf{R}^2$ para la cual

$$T(1, 0) = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

$$T(0, 1) = (a, b)$$

sea una transformación ortogonal.

7. Sea $\beta = \{v_1, v_2, v_3\}$ una base ortonormal de $\mathbf{R}^3$.
Determine una transformación ortogonal $T: \mathbf{R}^3 \to \mathbf{R}^3$ para la cual

$$T(v_1) = \frac{1}{3}(2v_1 + v_2 + 2v_3)$$

y

$$T(v_2) = \frac{1}{\sqrt{2}}(v_1 - v_2)$$

8. Se dice que la transformación lineal $T: V \to V$, en donde V es un espacio vectorial de dimensión finita con producto interno $(\cdot \mid \cdot)$, es una *transformación conforme*, si T preserva ángulos, es decir, si para cualesquiera dos vectores v y u de V el ángulo entre v y u es el mismo que el ángulo entre $T(v)$ y $T(u)$.
   a) Demuestre que si $T: V \to V$ es una transformación ortogonal entonces T es una transformación conforme.

   b) ¿Es válida la afirmación recíproca del inciso a)?
   (*Sugerencia*: considere $T: V \to V$, $T(v) = 2v$.)

9. Describa el efecto geométrico de la transformación $T: \mathbf{R}^2 \to \mathbf{R}^2$, $T(X) = AX$ en donde

   a) $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ c) $A = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$

   b) $A = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ d) $A = \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{bmatrix}$

10. Sean $T_1, T_2: V \to V$ dos transformaciones ortogonales tales que $det\ T_1 = det\ T_2 = 1$. Demuestre que $T_1 \circ T_2$ y $T_2 \circ T_1$ son también transformaciones ortogonales y que $det\ T_1 \circ T_2 = det\ T_2 \circ T_1 = 1$.

11. Interprete geométricamente el ejercicio anterior en el caso en el que V es el espacio vectorial $\mathbf{R}^2$ con el producto interno canónico.

② 12. ¿Es la suma de dos transformaciones ortogonales una transformación ortogonal? Si es así, demuéstrelo. Caso contrario, dé un contraejemplo.

13. Describa el efecto geométrico de la transformación $T: \mathbf{R}^2 \to \mathbf{R}^2$, $T(X) = AX$ en donde

   a) $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ c) $A = \begin{bmatrix} 1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}$

   b) $A = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix}$ d) $A = \begin{bmatrix} \sqrt{3}/2 & 1/2 \\ 1/2 & -\sqrt{3}/2 \end{bmatrix}$

③ 14. Si $T: V \to V$ es un operador lineal en el espacio vectorial V de dimensión 3, es posible demostrar (véase ejercicio 25 de la sección 1 del siguiente capítulo) que existe una base ortonormal β de V respecto de la cual la matriz que representa a T es de la forma

$$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & f & g \end{bmatrix}$$

Use este hecho para demostrar que la matriz que representa a una transformación ortogonal $T: \mathbf{R}^3 \to \mathbf{R}^3$ para la cual $det\ T = 1$ es de la forma

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\text{sen}\,\theta \\ 0 & \text{sen}\,\theta & \cos\theta \end{bmatrix}$$

para algún θ. ¿Cuál es el efecto geométrico de la transformación T?

④ 15. Sea A una matriz antisimétrica de orden n ($A^t = -A$)
   a) Demuestre que para cualquier vector $X \in \mathbf{R}^n$ se tiene $(AX \mid X) = 0$ (el producto interno canónico de $\mathbf{R}^n$).
   b) Compruebe que para $X \in \mathbf{R}^n$,

$$\|(A - I)X\|^2 = \|AX\|^2 + \|X\|^2$$

   c) Concluya del inciso anterior que la matriz $A - I$ es inversible.
   [*Sugerencia*: demuestre que $(A - I)\,X = 0 \Rightarrow X = 0$.]

d) Demuestre que para cualquier $X \in \mathbf{R}^n$ se tiene

$$\|(A + I)X\|^2 = \|(A - I)X\|^2$$

e) Sea $U = (A + I)(A - I)^{-1} \in M_{n \times n}$. Sea T: $\mathbf{R}^n \to \mathbf{R}^n$ el operador lineal $T(X) = UX$, pruebe que T es una isometría.

(*Sugerencia*: use el inciso anterior).

f) Demuestre que $U - I = 2\,(A - I)^{-1}$. Concluya entonces que $U - I$ es inversible.

g) Compruebe que si U es una matriz ortogonal (o bien, si el operador T: $\mathbf{R}^n \to \mathbf{R}^n$, $T(X) = UX$ es una isometría), tal que $U - I$ es inversible y si $A = (U + I)(U - I)^{-1}$, entonces A es una matriz antisimétrica.

(Nota: A la matriz U se le llama *transformada de Cayley* de la matriz A.)

APÉNDICE. FUNCIONALES LINEALES EN ESPACIOS CON PRODUCTO INTERNO

Sea $(V, (\cdot \mid \cdot))$ un espacio con producto interno.

Con el producto interno en V se pueden construir funcionales lineales f: $V \to \mathbf{R}$ de la siguiente manera:

Sea u_0 un vector fijo de V y defínase f: $V \to \mathbf{R}$ como

$$f(v) = (v \mid u_0)$$

El hecho de que f es lineal se sigue de las propiedades del producto interno:

$$f(v_1 + v_2) = (v_1 + v_2 \mid u_0) = (v_1 \mid u_0) + (v_2 \mid u_0) = f(v_1) + f(v_2) \qquad \forall\, v_1, v_2 \in V$$

$$f(cv) = (cv \mid u_0) = c(v \mid u_0) = cf(v) \qquad \forall\, c \in \mathbf{R}, \qquad \forall\, v \in V$$

Esta construcción da entonces una gran cantidad de ejemplos de funcionales lineales en el espacio V. Más aún, se demostrará que si V es de dimensión finita, esta "gran cantidad" abarca de hecho a *todos* los funcionales lineales en V; es decir, que para cualquier funcional lineal f en el espacio vectorial V de dimensión finita con producto interno $(\cdot \mid \cdot)$, existe un (único) vector $u_0 \in V$ tal que $f(v) = (v \mid u_0)\ \forall\, v \in V$.

TEOREMA A1

Sea V un espacio vectorial de dimensión finita con producto interno $(\cdot \mid \cdot)$, y sea f: $V \to \mathbf{R}$ una función lineal, entonces existe un único vector $u_0 \in V$ tal que

$$f(v) = (v \mid u_0) \qquad \forall\, v \in V$$

DEMOSTRACIÓN Se darán dos demostraciones distintas de la existencia del vector u_0.

DEMOSTRACIÓN 1 Según el corolario del teorema 4.2 existe una base ortonormal $\beta = \{v_1, v_2, \ldots, v_n\}$ de V.

Sea

$$u_0 = f(v_1)v_1 + f(v_2)v_2 + \ldots + f(v_n)v_n = \sum_{i=1}^{n} f(v_i)v_i$$

y sea g: $V \to \mathbf{R}$ el funcional lineal en V dado por

$$g(v) = (v \mid u_0)$$

Se afirma que $f = g$.

En efecto, para cualquier vector $v \in V$ se tiene

$$v = c_1v_1 + c_2v_2 + \ldots + c_nv_n = \sum_{j=1}^{n} c_jv_j$$

Entonces

$$g(v) = (v \mid u_0) = \left(\sum_{j=1}^{n} c_jv_j \,\middle|\, \sum_{i=1}^{n} f(v_i)v_i \right) = \sum_{j=1}^{n}\sum_{i=1}^{n} c_jf(v_i)(v_j \mid v_i)$$

$$= \sum_{j=1}^{n} c_jf(v_j) = f\left(\sum_{j=1}^{n} c_jv_j \right) = f(v)$$

lo que prueba que $g(v) = f(v)\ \ \forall\, v \in V$, es decir, que $g = f$.

DEMOSTRACIÓN 2 Considérese el núcleo de f

$$Ker\,f = \{v \in V \mid f(v) = 0\}$$

Se sabe que $Ker\,f$ es un subespacio de V. Si $Ker\,f = V$, entonces $f(v) = 0$, $\forall\, v \in V$, y en tal caso, el teorema se cumple con $u_0 = 0$ (esto es, $f(v) = (v \mid 0)\ \forall\, v \in V$).

Considérese entonces el caso en el que $Ker\,f \neq V$.

En tal caso, el subespacio $(Ker\,f)^{\perp}$ es distinto del subespacio trivial $\{0\}$ (de hecho $dim\,(Ker\,f)^{\perp} = 1$, ¿por qué?). Existe entonces un vector $w \in (Ker\,f)^{\perp}$, $w \neq 0$.

Sea $v \in V$ y considere el vector

$$u = f(v)w - f(w)v$$

Obsérvese que

$$f(u) = f(f(v)w - f(w)v) = f(v)f(w) - f(w)f(v) = 0$$

por lo que $u \in Ker\, f$. Entonces $(u \mid w) = 0$. Es decir,

$$(f(v)w - f(w)v \mid w) = 0$$

expresión que puede reescribirse como

$$f(v)(w \mid w) = f(w)(v \mid w)$$

o bien,

$$f(v) = \left(v \;\middle|\; \frac{f(w)}{\|w\|^2}\, w\right)$$

Defínase

$$u_0 = \frac{f(w)}{\|w\|^2}\, w$$

(obsérvese que $u_0 \in (Ker\, f)^\perp$). Se ha probado entonces que para cualquier otro vector $v \in V$ se tiene

$$f(v) = (v \mid u_0)$$

Para finalizar la demostración del teorema, se probará que el vector u_0 es único.

Supóngase por contradicción que existe otro vector $u_1 \in V$ tal que

$$f(v) = (v \mid u_1) \qquad \forall\, v \in V$$

Entonces

$$(v \mid u_0) = (v \mid u_1) \qquad \forall\, v \in V$$

o bien,

$$(v \mid u_0 - u_1) = 0 \qquad \forall\, v \in V$$

En particular, si $v = u_0 - u_1$ se obtiene

$$(u_0 - u_1 \mid u_0 - u_1) = 0$$

de donde $u_0 - u_1 = 0$, o sea $u_0 = u_1$.

Q.E.D.

El resultado anterior es en general falso cuando el espacio vectorial V no es de dimensión finita.

Por ejemplo, considérese el espacio $C([0, 1])$ de las funciones continuas en el intervalo $[0, 1]$ con el producto interno

$$(f \mid g) = \int_0^1 f(x)g(x)dx$$

Se verá que existe (al menos) un funcional lineal F: $C([0, 1]) \to \mathbf{R}$ que no puede ser representado como un producto interno en $C([0, 1])$. Es decir, que no existe $g \in C([0, 1])$ tal que

$$F(f) = (f \mid g) \qquad \forall\, f \in C[0, 1])$$

Sea F: $C([0, 1]) \to \mathbf{R}$ el funcional lineal dado por

$$F(f) = f(1/2) \qquad f \in C([0, 1])$$

y supóngase, por contradicción, que existe $g \in C([0, 1])$ tal que

$$F(f) = (f \mid g) = \int_0^1 f(x)g(x)dx = f(1/2) \qquad \forall\, f \in C([0, 1])$$

En particular, si $f(x) = (x - 1/2)^2 g(x) \in C([0, 1])$ se tendría

$$(f \mid g) = \int_0^1 (x - 1/2)^2 g(x)g(x)dx = (1/2 - 1/2)^2 g(1/2) = 0$$

Es decir,

$$\int_0^1 (x - 1/2)^2 (g(x))^2 = 0 \qquad \forall\, x \in [0, 1]$$

Como la función en el integrando de la expresión anterior es una función continua en $[0, 1]$ *no negativa*, se concluye que

$$(x - 1/2)^2 (g(x))^2 = 0 \qquad \forall\, x \in [0, 1]$$

Entonces $g(x) = 0\ \forall\, x \in [0, 1]$ (¿por qué?). Es decir, $g = 0$, lo cual implica que $F(f) = (f \mid g) = (f \mid 0) = 0\ \forall\, f \in C([0, 1])$, que es una contradicción.

Nota: Cuando se consideran funcionales lineales en espacios vectoriales $(V, (\cdot \mid \cdot))$ de dimensión infinita con producto interno, es posible demostrar la validez de un resultado análogo al presentado en el teorema A1, conocido como "teorema de representación (de funcionales lineales acotados) de Riesz-Fisher", si además se tienen un par de hipótesis adicionales:

1) El funcional lineal f: $V \to \mathbf{R}$ es *acotado*. Esto significa que existe una constante $K \geq 0$ tal que

$$|f(v)| \leq K\|v\| \qquad \forall\, v \in V$$

2) El espacio vectorial $(V, (\cdot|\cdot))$ es completo respecto de la métrica

$$d(v_1, v_2) = \|v_1 - v_2\|$$

en donde $\|\cdot\| = \sqrt{(\cdot|\cdot)}$. Esto significa que siempre que se tenga una sucesión (v_n) en V para la cual dado $\epsilon > 0$ existe $n_0 \in N$ tal que

$$d(v_n, v_m) < \epsilon \qquad \forall\, n, m \geq N$$

(una sucesión de este tipo es llamada sucesión de Cauchy), siempre existe un vector $v_0 \in V$ tal que

$$\lim v_n = v_0$$

Es decir $(V, (\cdot|\cdot))$ es un espacio en el cual toda sucesión de Cauchy tiene límite. A los espacios vectoriales con producto interno con esta propiedad se les llama *espacios de Hilbert.*

Resulta que cuando $(V, (\cdot|\cdot))$ es un espacio vectorial de dimensión finita con producto interno, las dos hipótesis anteriores se satisfacen automáticamente. El lector interesado puede consultar el capítulo 4 del libro de W. Rudin, "*Real and Complex Analysis*" (3a. edición), en donde se presenta una sencilla introducción al estudio de los espacios de Hilbert.

EJERCICIOS (APÉNDICE, CAPÍTULO 5)

1. Compare la demostración 1 del teorema A1 con el teorema A_I1 del apéndice I del capítulo 4.
2. Sea $(V, (\cdot|\cdot))$ un espacio vectorial de dimensión finita con producto interno. Considere el espacio dual V^* de V (véase apéndice I del capítulo 4). Dado $f \in V^*$ se ha visto que existe un único vector $u_f \in V$ tal que

$$f(v) = (v|u_f) \qquad v \in V$$

Considere la transformación $T: V^* \to V$ dada por $T(f) = u_f$.

Demuestre que T es un isomorfismo.
3. Sea v_0 un vector del espacio vectorial $(V, (\cdot|\cdot))$ de dimensión finita con producto interno. Si $f \in V^*$, dé sentido, en base al ejercicio anterior, a la siguiente definición: "se dice que el funcional $f \in V^*$ es ortogonal a v_0 si $f(v_0) = 0$".
4. Con la definición del ejercicio anterior, demuestre que si el espacio vectorial de dimensión finita V es la suma directa de los subespacios U y W, entonces su espacio dual V^* es la suma directa de los complementos ortogonales de U y W, $U^\perp$ y $W^\perp$.

5. Sea $(V, (\cdot|\cdot))$ un espacio vectorial de dimensión finita con producto interno, si H es un hiperespacio de V (véase sección 3 del apéndice I. Capítulo 4), pruebe que

$$H = \{v \in V \mid (v|v_0) = 0\}$$

para algún $v_0 \in V$.
6. Sea $(V, (\cdot|\cdot))$ un espacio vectorial de dimensión finita con producto interno. Si M es un hiperplano de V, demuestre que

$$M = \{v \in V \mid (v|v_0) = c\}$$

para algún $v_0 \in V$ y algún escalar c.
7. Considere el espacio $\mathbf{R}^3$ con el producto interno canónico. Demuestre que las siguientes afirmaciones acerca de un subconjunto M de $\mathbf{R}^3$ son equivalentes.
   a) M es un plano que pasa por el origen.
   b) M es un subespacio de $\mathbf{R}^3$ de dimensión 2.
   c) M puede escribirse como

$$M = \{(x, y, z) \in \mathbf{R}^3 \mid ax + by + cz = 0\}$$

   para ciertos escalares a, b y c no todos nulos.
   d) M puede escribirse como

$$M = \{v \in \mathbf{R}^3 \mid f(v) = 0\}$$

   para algún funcional lineal no nulo $f \in (\mathbf{R}^3)^*$.
   e) M puede escribirse como

$$M = \{v \in \mathbf{R}^3 \mid (v|v_0) = 0\}$$

   para algún vector $v_0 \in \mathbf{R}^3$ no nulo.

CAPÍTULO SEIS

Valores y vectores propios

Considérese el operador lineal $T: \mathbf{R}^3 \to \mathbf{R}^3$ dado por:

$$T(x, y, z) = (x - z, x + 2y + z, 2x + 2y + 3z)$$

La matriz que representa a T respecto de la **base canónica** β_1 de $\mathbf{R}^3$

$$A = [T]_{\beta_1} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix}$$

Véase cuál es la representación matricial del operador T respecto de la base $\beta_2 = \{v_1, v_2, v_3\}$ en donde:

$$v_1 = (1, -1, 0)$$
$$v_2 = (2, -1, -2)$$
$$v_3 = (1, -1, -2)$$

Para un vector arbitrario $(x_0, y_0, z_0) \in \mathbf{R}^3$ se tiene que:

$$((x_0, y_0, z_0))_{\beta_2} = \left(-y_0 + \frac{1}{2}z_0, x_0 + y_0, -\frac{1}{2}z_0 - x_0 - y_0\right)$$

Entonces,

$$T(v_1) = T(1, -1, 0) = (1, -1, 0) \qquad ((1, -1, 0))_{\beta_2} = (1, 0, 0)$$
$$T(v_2) = T(2, -1, -2) = (4, -2, -4) \qquad ((4, -2, -4))_{\beta_2} = (0, 2, 0)$$
$$T(v_3) = T(1, -1, -2) = (3, -3, -6) \qquad ((3, -3, -6))_{\beta_2} = (0, 0, 3)$$

de modo que la matriz de T respecto de la base β_2 es:

$$B = [T]_{\beta_2}^{\beta_2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

De la matriz diagonal B que representa al operador T respecto de la base β_2 se puede obtener "a simple vista" la siguiente información acerca de T

1) El rango de T es 3 (pues el rango de B es 3), y por lo tanto, la nulidad de T es 0.
2) El determinante de T es 6.
3) T es un operador inversible.
4) La matriz que representa al operador inverso de T, $T^{-1}: \mathbf{R}^3 \to \mathbf{R}^3$ respecto de la base β_2 es:

$$B^{-1} = [T^{-1}]_{\beta_2} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}$$

Ciertamente esta información se podría haber obtenido también de la representación matricial de T dada por la matriz A. Sin embargo, debe resultar claro que la cantidad de trabajo que se requiere para llegar a establecer, por ejemplo, la inversibilidad de T a partir de la matriz A, es mucho mayor respecto del trabajo requerido para concluir la inversibilidad de T a partir de la matriz B.

Obsérvese que los vectores de la base β_2, respecto de la cual T tiene una representación matricial agradable, tienen la particularidad siguiente:

$$T(v_1) = T((1, -1, 0)) = (1, -1, 0) = v_1$$
$$T(v_2) = T((2, -1, -2)) = (4, -2, -4) = 2(2, -1, -2) = 2v_2$$
$$T(v_3) = T((1, -1, -2)) = (3, -3, -6) = 3(1, -1, -2) = 3v_3$$

Es decir, las imágenes de cada uno de los vectores de la base β_2 resultan ser múltiplos de ellos mismos.

En general, uno se podría preguntar si siempre es posible conseguir una base del espacio vectorial V, tal que la representación matricial del operador $T: V \to V$ sea una matriz diagonal. Esta misma pregunta se puede plantear en términos "puramente matriciales": bajo la relación de semejanza en el conjunto de matrices cuadradas (véase sección 4 del capítulo 4), las matrices A y B del ejemplo anterior pertenecen a la misma clase de equivalencia (esto es, A y B son semejantes). Se ha visto entonces que en la clase de equivalencia de la matriz A existe una matriz diagonal (la matriz B). En general, uno se pregunta si dada una matriz cuadrada A, existe una matriz diagonal B semejante a A, esto es, una matriz diagonal B en la clase de equivalencia de A.

Uno de los objetivos principales de este capítulo es dar respuesta a esta pregunta (la respuesta será —desgraciadamente— *NO*).

El primer paso que se dará es estudiar lo relacionado con vectores $v \in V$ cuyas imágenes bajo la transformación lineal $T: V \to V$ sean múltiplos de ellos mismos (en el ejemplo analizado anteriormente, ésta es la característica de los vectores de la base β_2 respecto de la cual la matriz que representa a T es diagonal).

1. DEFINICIÓN Y RESULTADOS PRELIMINARES

DEFINICIÓN 1.1.a) Sea A una matriz cuadrada de orden n. El número λ es llamado *valor propio* (valor característico o eigenvalor) de A si existe un vector $X \in \mathbf{R}^n$ no nulo tal que $AX = \lambda X$. En tal caso, al vector X se le llama *vector propio* (vector característico o eigenvector) correspondiente (o asociado) al valor propio λ.

DEFINICIÓN 1.1.b) Sea $T: V \to V$ un operador lineal en el espacio vectorial V de dimensión n. El número λ es llamado *valor propio* (valor característico o eigenvalor) de T si existe un vector $v \in V$ no nulo, tal que $T(v) = \lambda v$. En tal caso, al vector v se le llama *vector propio* (vector característico o eigenvector) correspondiente (o asociado) al valor propio λ.

Existe una equivalencia natural entre las definiciones 1.1. a) y 1.1. b), proporcionada por la correspondencia matrices —transformaciones lineales, según se muestra a continuación—: sea V un espacio vectorial de dimensión n y sea $T: V \to V$ un operador lineal en V. Si $\beta = \{v_1, v_2, \ldots, v_n\}$ es una base de V, considérese la matriz A de T respecto de la base β. Se tiene entonces que

$$\lambda \text{ es un valor propio de } T \Leftrightarrow \lambda \text{ es un valor propio de } A = [T]_\beta$$

En efecto, supóngase que λ es un valor propio de T y que $v \in V$ es un vector propio asociado a λ. Entonces, según la definición 1.1. b), se tiene

$$T(v) = \lambda v$$

de donde

$$[T(v)]_\beta = [\lambda v]_\beta$$

o

$$[T]_\beta[v]_\beta = \lambda[v]_\beta$$

es decir,

$$AX = \lambda X$$

en donde $X = [v]_\beta$. Por lo tanto, λ es un valor propio de A y $X = [v]_\beta$ es un vector propio asociado a λ (obsérvese que $v \neq 0 \Rightarrow X \neq 0$).

Recíprocamente, supóngase que λ es un valor propio de $A = [T]_\beta$ y que $X = (c_1, c_2, \ldots, c_n) \in \mathbf{R}^n$ es un vector propio asociado a λ. Sea $v = c_1v_1 + c_2v_2 + \ldots + c_nv_n \in V$ (esto es, $X = (v)_\beta$), entonces, según la definición 1.1. a), se tiene

$$AX = \lambda X$$

es decir,

$$[T]_\beta[v]_\beta = \lambda[v]_\beta$$

lo cual se puede escribir como

$$[T(v)]_\beta = [\lambda v]_\beta$$

de donde

$$T(v) = \lambda v$$

lo que muestra que λ es un valor propio de T y que $v \in V$ (que es no nulo pues $X = (v)_\beta$ es no nulo) es un vector propio asociado a λ.

Más aún, supóngase que se toma otra base β' de V distinta de la base β. Sea $B = [T]_{\beta'}$. Se sabe que

$$A = P^{-1}BP$$

para alguna matriz inversible P. Si λ es un valor propio de A, entonces existe $X \in \mathbf{R}^n$ no nulo, tal que

$$AX = \lambda X$$

o bien,

$$(P^{-1}BP)X = \lambda X$$

expresión que se puede reescribir como (multiplicando por la izquierda por la matriz P y asociando)

$$B(PX) = \lambda(PX)$$

Como $X \neq 0$ y P es una matriz inversible, entonces $PX \neq 0$ (¿por qué?) de manera que λ es también un valor propio de la matriz B (y PX es un vector propio asociado a λ).

Obsérvese entonces que se puede replantear la definición 1.1. b), como:

DEFINICIÓN 1.1.b)' Sea $T: V \to V$ un operador lineal en el espacio vectorial V de dimensión n. El número λ es llamado valor propio de T si λ es un valor propio de la matriz A asociada al operador T respecto de alguna (de cualquiera) base de V.

EJEMPLO 1 Por ejemplo, sea A la matriz

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$$

El número $\lambda = 5$ es un valor propio de A pues para el vector no nulo $X = (1, 3)$ en $\mathbf{R}^2$ se tiene

$$AX = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 15 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 3 \end{bmatrix} = 5X$$

Si se considera al operador $T: \mathbf{R}^2 \to \mathbf{R}^2$ dado por $T(X) = AX$, entonces, según el análisis anterior, $\lambda = 5$ será también un valor propio de T. Más aún, observe que T puede describirse explícitamente como

$$T(x, y) = (2x + y, 3x + 4y)$$

La representación matricial de T respecto de la base $\beta = \{(1, 1), (0, 1)\}$ está dada por la matriz

$$B = \begin{bmatrix} 3 & 1 \\ 4 & 3 \end{bmatrix}$$

Según se vio anteriormente, $\lambda = 5$ deberá ser también un valor propio de la matriz B. En efecto, para el vector $X = (1, 2)$ se tiene

$$BX = \begin{bmatrix} 3 & 1 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 5X$$

1.1. VALORES PROPIOS

El siguiente teorema da luz acerca de la manera como se pueden determinar los valores propios de una matriz.

TEOREMA 1.1 Sea A una matriz cuadrada, el número λ es un valor propio de A si, y sólo si $det(A - \lambda I) = 0$.

DEMOSTRACIÓN Si λ es un valor propio de A, existe un vector $X \in \mathbf{R}^n$ no nulo tal que $AX = \lambda X$. Esta expresión se puede reescribir como $(A - \lambda I)X = 0$. Se puede contemplar esta última expresión como un sistema homogéneo de ecuaciones lineales (la matriz del sistema siendo $A - \lambda I$) el cual se sabe que posee soluciones no triviales (. . . existe un vector X no nulo . . .). Esto equivale, como se sabe ya, a que la matriz del sistema $A - \lambda I$ sea no inversible. En tal caso, $det(A - \lambda I) = 0$.

Q.E.D.

Obsérvese que si $A = (a_{ij})_{i,j = 1,2,\ldots,n}$ entonces,

$$A - \lambda I = \begin{bmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{bmatrix}$$

Uno de los productos elementales que aparecen en el desarrollo del determinante de esta matriz es el producto de los elementos de su diagonal principal

$$(a_{11} - \lambda)(a_{22} - \lambda) \ldots (a_{nn} - \lambda)$$

Aún mas, en todos los restantes productos elementales del desarrollo de $det(A - \lambda I)$ aparecerá siempre como factor al menos un elemento fuera de la diagonal principal de $A - \lambda I$. Entonces, al desarrollar $det(A - \lambda I)$ se obtendrá un *polinomio en λ de grado n* (de hecho, el término $\alpha\lambda^n$ de este polinomio proviene del producto elemental arriba mencionado, con $\alpha = (-1)^n$).

Por ejemplo, si A es la matriz

$$A = \begin{bmatrix} 3 & -1 \\ 2 & 4 \end{bmatrix}$$

se tiene

$$det(A - \lambda I) = det\begin{bmatrix} 3 - \lambda & -1 \\ 2 & 4 - \lambda \end{bmatrix} = (3 - \lambda)(4 - \lambda) + 2 = \lambda^2 - 7\lambda + 14$$

y si

$$B = \begin{bmatrix} 3 & 2 & 8 \\ 0 & 5 & 7 \\ 0 & 0 & 7 \end{bmatrix}$$

se tiene

$$det(B - \lambda I) = \begin{bmatrix} 3 - \lambda & 2 & 8 \\ 0 & 5 - \lambda & 7 \\ 0 & 0 & 7 - \lambda \end{bmatrix} = (3 - \lambda)(5 - \lambda)(7 - \lambda)$$

$$= -\lambda^3 + 15\lambda^2 - 71\lambda + 105$$

DEFINICIÓN 1.2 Sea A una matriz cuadrada de orden n, al polinomio (de grado n en λ) $p(\lambda) = det(A - \lambda I)$ se le llama *polinomio característico* de A.

Obsérvese entonces que el teorema 1.1 establece que λ es un valor propio de la matriz A si, y sólo si $p(\lambda) = 0$, o sea, si, y sólo si λ *es una raíz del polinomio característico de A*. Este hecho da entonces la pauta para encontrar los valores propios de una matriz A: se halla primeramente el polinomio característico de A y se determinan las raíces de éste. Éstas serán los valores propios de A.

Véanse algunos ejemplos.

EJEMPLO 2 Si A es la matriz

$$A = \begin{bmatrix} 5 & 9 \\ 2 & 2 \end{bmatrix}$$

se tiene

$$p(\lambda) = det(A - \lambda I) = det\begin{bmatrix} 5-\lambda & 9 \\ 2 & 2-\lambda \end{bmatrix} = (5-\lambda)(2-\lambda) - 18$$
$$= \lambda^2 - 7\lambda - 8 = (\lambda - 8)(\lambda + 1)$$

Las raíces de $p(\lambda)$ son entonces, $\lambda_1 = 8$ y $\lambda_2 = -1$. Éstos son los valores propios de A.

EJEMPLO 3 Para la matriz

$$B = \begin{bmatrix} 1 & 0 \\ 8 & 1 \end{bmatrix}$$

se tiene

$$p(\lambda) = det(B - \lambda I) = det\begin{bmatrix} 1-\lambda & 0 \\ 8 & 1-\lambda \end{bmatrix} = (1-\lambda)^2$$

Entonces, el único valor propio de B es $\lambda = 1$.

EJEMPLO 4 Por último, para la matriz

$$C = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}$$

se tiene

$$p(\lambda) = det(C - \lambda I) = det\begin{bmatrix} 2-\lambda & -1 \\ 3 & 2-\lambda \end{bmatrix} = (2-\lambda)^2 + 3 = \lambda^2 - 4\lambda + 7$$

La ecuación $p(\lambda) = \lambda^2 - 4\lambda + 7 = 0$ tiene por raíces

$$\frac{4 \pm \sqrt{16-28}}{2} = \frac{4 \pm \sqrt{-12}}{2} = \frac{4 \pm 2\sqrt{3}\,i}{2} = 2 \pm \sqrt{3}\,i$$

Los números con los que se está trabajando en este libro son números reales (en los espacios vectoriales que ya han sido considerados se ha restringido al campo de los números reales). Entonces, para la matriz C del ejemplo anterior se dirá que "no tiene valores propios" significando con esto que la matriz C no tiene valores propios *reales*.

1.2. VECTORES PROPIOS. ESPACIOS PROPIOS

Se ha visto ya cómo se pueden determinar los valores propios de una matriz cuadrada A. Se verá ahora cómo determinar los vectores propios correspondientes.

Sea λ un valor propio de A. Entonces, $det(A - \lambda I) = 0$. El vector $X \in \mathbf{R}^n$ es un vector propio de A asociado al valor propio λ si X es no nulo y si $AX = \lambda X$. En otras palabras, X es un vector propio asociado al valor propio λ si X es una solución no trivial de $(A - \lambda I)X = 0$.

La existencia de soluciones no triviales de este sistema homogéneo de ecuaciones lineales (esto es, la existencia de vectores propios de A asociados al valor propio λ) está asegurada por el hecho de que λ siendo un valor propio de A, la matriz $A - \lambda I$ es no inversible [pues $det(A - \lambda I) = 0$].

En resumen, para determinar vectores propios asociados al valor propio λ, sólo se tienen que hallar soluciones no triviales del sistema homogéneo de ecuaciones lineales $(A - \lambda I)X = 0$.

EJEMPLO 5 Por ejemplo, se había visto ya que los valores propios de la matriz

$$A = \begin{bmatrix} 5 & 9 \\ 2 & 2 \end{bmatrix}$$

son $\lambda_1 = 8$ y $\lambda_2 = -1$. Para determinar los vectores propios correspondientes a cada uno de estos valores propios, se tiene que considerar el sistema homogéneo $(A - \lambda I)X = 0$; es decir,

$$\begin{bmatrix} 5-\lambda & 9 \\ 2 & 2-\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Para el valor $\lambda_1 = 8$ se tiene el sistema

$$\begin{bmatrix} -3 & 9 \\ 2 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = 3r$, $x_2 = r$, $r \in \mathbf{R}$. Entonces, los vectores propios de A asociados al valor propio $\lambda_1 = 8$ son vectores de la forma $(3r, r)$, $r \in \mathbf{R}$, $(r \neq 0)$.

Para el valor propio $\lambda_2 = -1$ el sistema por resolver es:

$$\begin{bmatrix} 6 & 9 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = -\frac{3}{2}r$, $x_2 = r$, $r \in \mathbf{R}$. Entonces, los vectores propios de A asociados al valor propio $\lambda_2 = -1$ son de la forma $\left(-\frac{3}{2}r, r\right)$, $r \in \mathbf{R}$, $(r \neq 0)$.

Obsérvese que el conjunto de vectores propios de una matriz A asociados al valor propio λ, siendo éste el conjunto de soluciones *no triviales* del sistema

homogéneo $(A - \lambda I)X = 0$, no es en sí mismo un subespacio de $\mathbf{R}^n$ (pues carece del vector cero). Sin embargo, si se considera la unión de este conjunto con el (conjunto formado por) el vector cero, lo que se obtiene es precisamente el espacio solución del sistema homogéneo $(A - \lambda I)X = 0$, el cual se sabe que es un subespacio de $\mathbf{R}^n$

DEFINICIÓN 1.3 Sea A una matriz cuadrada de orden n y sea λ un valor propio de A. Sea E_λ la unión del conjunto de vectores propios de A asociados al valor propio λ con (el conjunto formado por) el vector cero. E_λ es entonces un subespacio de $\mathbf{R}^n$ llamado *espacio propio* (espacio característico o eigenespacio) de la matriz A correspondiente al valor propio λ.

EJEMPLO 6 En el ejemplo visto anteriormente, para la matriz

$$A = \begin{bmatrix} 5 & 9 \\ 2 & 2 \end{bmatrix}$$

se tiene

$$E_{\lambda_1 = 8} = \{(3t, t), t \in \mathbf{R}\}$$

$$E_{\lambda_2 = -1} = \left\{\left(-\frac{3}{3}t, t\right), t \in \mathbf{R}\right\}$$

Véase un ejemplo más.

EJEMPLO 7 Considere la matriz

$$A = \begin{bmatrix} 1 & 3 & 0 \\ 0 & -2 & 0 \\ 0 & 6 & 1 \end{bmatrix}$$

El polinomio característico de A es:

$$p(\lambda) = det(A - \lambda I) = det\begin{bmatrix} 1-\lambda & 3 & 0 \\ 0 & -2-\lambda & 0 \\ 0 & 6 & 1-\lambda \end{bmatrix}$$

Las raíces del polinomio característico A son $\lambda_1 = 1$ y $\lambda_2 = -2$. Éstos son los valores propios de A. Determine los espacios propios correspondientes a cada uno de estos valores propios.

Se tiene que considerar el sistema homogéneo $(A - \lambda I)X = 0$. Es decir,

$$\begin{bmatrix} 1-\lambda & 3 & 0 \\ 0 & -2-\lambda & 0 \\ 0 & 6 & 1-\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Para $\lambda_1 = 1$ se tiene el sistema

$$\begin{bmatrix} 0 & 3 & 0 \\ 0 & -3 & 0 \\ 0 & 6 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = r$, $x_2 = 0$, $x_3 = s$, $r, s \in \mathbf{R}$. Entonces, el espacio propio E_{λ_1} es:

$$E_{\lambda_1} = (r, 0, s), \quad r, s \in \mathbf{R}$$

la cual tiene dimensión 2 [una base de él está constituida por los vectores (1, 0, 0) y (0, 0, 1)].

Para $\lambda_2 = -2$ se tiene el sistema

$$\begin{bmatrix} 3 & 3 & 0 \\ 0 & 0 & 0 \\ 0 & 6 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = \frac{1}{2}r$, $x_2 = -\frac{1}{2}r$, $x_3 = r$, $r \in \mathbf{R}$. Entonces, el espacio propio E_{λ_2} es:

$$E_{\lambda_2} = \left\{\left(\frac{1}{2}r, -\frac{1}{2}r, r\right), \quad r \in \mathbf{R}\right\}$$

el cual tiene por dimensión 1 [una base de él es $\{(1, -1, 2)\}$].

Hasta este momento, se ha estudiado lo relacionado con valores, vectores y espacios propios para una matriz cuadrada A de orden n. Véase ahora cómo se ven estos mismos conceptos cuando se consideran operadores lineales en lugar de matrices.

El siguiente teorema es el análogo del teorema 1.1. En él se da una caracterización de los valores propios de un operador en términos de la inversibilidad del operador $T - \lambda I$, en donde I denota al operador identidad $I: V \to V$, $I(v) = v$, $v \in V$.

TEOREMA 1.2 Sea $T: V \to V$ un operador lineal en el espacio vectorial V de dimensión n, el número λ es un valor propio de T si, y sólo si el operador $T - \lambda I$: $V \to V$ es no inversible.

DEMOSTRACIÓN Se tienen las siguientes equivalencias (véase teorema 6.4 del capítulo 4).

λ es un valor propio de $T \Leftrightarrow$ existe $v \in V$, $v \neq 0$, tal que $T(v) = \lambda v \Leftrightarrow$ existe $v \in V$, $v \neq 0$, tal que $(T - \lambda I)(v) = 0 \Leftrightarrow Ker\,(T - \lambda I) \neq \{0\} \Leftrightarrow T - \lambda I$ no es inversible.

Q.E.D.

Se define ahora el polinomio característico de un operador lineal $T: V \to V$. Para poder hacerlo, se necesitará el siguiente teorema:

TEOREMA 1.3

> Si A y B son matrices semejantes, entonces el polinomio característico de A coincide con le polinomio característico de B. En particular, A y B tienen los mismos valores propios.

DEMOSTRACIÓN Si A y B son matrices semejantes, entonces existe una matriz inversible P tal que $A = P^{-1}BP$. En tal caso,

$$\begin{aligned} det(A - \lambda I) &= det(P^{-1}BP - \lambda I) = det(P^{-1}(B - \lambda I)P) \\ &= (det\, P^{-1})(det(B - \lambda I))(det\, P) \\ &= (det\, P)^{-1}(det(B - \lambda I))(det\, P) = det(B - \lambda I) \end{aligned}$$

Entonces, el polinomio característico de A ($= det(A - \lambda I)$) y el polinomio característico de B ($= det(B - \lambda I)$) coinciden.

Q.E.D.

Se puede establecer entonces la siguiente definición:

DEFINICIÓN 1.4 Sea $T: V \to V$ un operador lineal en el espacio vectorial V de dimensión n. Se llama polinomio característico del operador T al polinomio característico de la matriz A que representa a T en alguna (cualquiera) base de V.

La definición anterior hace perfecto sentido, pues si β y β' son dos bases distintas de V, las matrices $A = [T]_\beta$ y $B = [T]_{\beta'}$, son semejantes. En tal caso, según el teorema 1.3, el polinomio característico de A es igual al polinomio característico de B (igual al polinomio característico de T).

Obsérvese por último, que si el vector $v \in V$ es un vector propio de T asociado al valor propio λ, entonces $v \neq 0$ y $T(v) = \lambda v$, o bien, v es un vector no nulo tal que $(T - \lambda I)v = 0$. Si se considera la unión del conjunto de vectores propios de T con (el conjunto formado por) el vector cero, se obtiene así el conjunto de todos los vectores $v \in V$ tales que $(T - \lambda I)v = 0$. Estos vectores son precisamente los que constituyen el núcleo del operador $T - \lambda I$, el cual se sabe que es un subespacio de V. Éste es entonces el espacio propio E_λ asociado al valor propio λ. Es decir,

$$E_\lambda = Ker\,(T - \lambda I)$$

Para finalizar esta sección, se verá un teorema que relaciona la dimensión del espacio propio E_λ correspondiente al valor propio λ de una matriz A (o de un operador T) con la multiplicidad de λ como raíz del polinomio característico de A (o de T).

TEOREMA 1.4

> Sea A una matriz cuadrada de orden n y sea λ un valor propio de A. Si E_λ es el espacio propio asociado al valor propio λ, entonces la dimensión de E_λ no excede a la multiplicidad de λ como raíz del polinomio característico $p(\lambda) = det\,(A - \lambda I)$.

DEMOSTRACIÓN Sea $T: \mathbf{R}^n \to \mathbf{R}^n$ el operador lineal $T(X) = AX$. Sea $\beta_1 = \{v_1, v_2, \ldots, v_k\}$ una base del espacio propio E_λ, existen vectores $v_{k+1}, \ldots, v_n$ tales que $\beta = \{v_1, v_2, \ldots, v_n\}$ es una base de $\mathbf{R}^n$. Sea $B = [T]_\beta$. En la j-ésima columna de B aparecen las coordenadas del vector $(T(v_j))_\beta$.

Obsérvese que para $1 \leq j \leq k$ se tiene

$$T(v_j) = \lambda v_j$$

pues los vectores $v_1, v_2, \ldots, v_k$ de la base β_1 de E_λ son —por definición— vectores propios de A asociados al valor propio λ. Entonces, como $(T(v_j))_\beta = (\lambda(v_j))_\beta = (0, 0, \ldots, \lambda, \ldots, 0)$ se tiene que B es una matriz de la forma

$$B = \left[\begin{array}{ccc|c} \lambda & & & D \\ & \ddots & & \\ & & \lambda & \\ \hline & 0 & & C \end{array}\right] \quad k$$

en donde C es una matriz cuadrada de orden $n - k$ y D es una matriz de orden $k \times (n - k)$.

Las matrices A y B son semejantes (ambas representan al mismo operador T) y por lo tanto, tienen el mismo polinomio característico $p(x)$. Este polinomio es:

$$p(x) = det(B - xI) = det \left[\begin{array}{ccc|c} \lambda - x & & & D \\ & \ddots & & \\ & & \lambda - x & \\ \hline & 0 & & C - xI_{n-k} \end{array}\right]$$

Al desarrollar por cofactores a lo largo de la primera columna y repitiendo k veces este desarrollo en los determinantes obtenidos, se llega finalmente a:

$$p(x) = (\lambda - x)^k\ det(C - xI_{n-k})$$

Entonces, la multiplicidad de λ como raíz del polinomio característico de A es al menos $k = dim\ E_\lambda$. Esto era precisamente lo que se quería probar.

Q.E.D.

EJEMPLO 8 Como un último ejemplo considere la matriz

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 3 & 3 & 1 \end{bmatrix}$$

El polinomio característico de A es:

$$p(\lambda) = det(A - \lambda I) = det\begin{bmatrix} 1-\lambda & 1 & 1 & 1 \\ 0 & 2-\lambda & 0 & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 3 & 3 & 1-\lambda \end{bmatrix}$$

$$= (1-\lambda)\, det\begin{bmatrix} 2-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 3 & 3 & 1-\lambda \end{bmatrix}$$

$$= (1-\lambda)(2-\lambda)\, det\begin{bmatrix} 1-\lambda & 0 \\ 3 & 1-\lambda \end{bmatrix}$$

$$= (1-\lambda)(2-\lambda)(1-\lambda)^2 = (1-\lambda)^3(2-\lambda)$$

Por lo tanto, los valores propios de A son $\lambda_1 = 1$ (raíz de multiplicidad 3 en el polinomio $p(\lambda)$) y $\lambda_2 = 2$ (raíz de multiplicidad 1 en $p(\lambda)$).

Para hallar los vectores propios correspondientes se tiene que considerar el sistema homogéneo:

$$\begin{bmatrix} 1-\lambda & 1 & 1 & 1 \\ 0 & 2-\lambda & 0 & 0 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 3 & 3 & 1-\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Para el valor propio $\lambda_1 = 1$, este sistema es:

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 3 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = r$, $x_2 = x_3 = x_4 = 0$, $r \in \mathbf{R}$. Entonces el espacio propio E_{λ_1} es:

$$E_{\lambda_1} = \{(r, 0, 0, 0), r \in \mathbf{R}\}$$

La dimensión de E_{λ_2} es entonces 1 (una base de él es $\{(1, 0, 0, 0)\}$). Se tiene, tal

como lo asegura el teorema 1.4

$$dim\ E_{\lambda_1} = 1 \leq 3 = \text{multiplicidad de } \lambda_1 \text{ en } p(\lambda).$$

Para el valor propio $\lambda_2 = 2$, el sistema homogéneo por resolver es:

$$\begin{bmatrix} -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 3 & 3 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = \frac{4}{3}r$, $x_2 = \frac{1}{3}r$, $x_3 = 0$, $x_4 = r$, $r \in \mathbf{R}$. Entonces, el espacio propio E_{λ_2} es:

$$E_{\lambda_2} = \left\{\left(\frac{4}{3}r, \frac{1}{3}r, 0, r\right), r \in \mathbf{R}\right\}$$

el cual tiene dimensión 1 (una base de él es $\{(4, 1, 0, 3)\}$).

De nueva cuenta se tiene, tal como lo asegura el teorema 1.4,

$$dim\ E_{\lambda_2} = 1 \leq 1 = \text{multiplicidad de } \lambda_2 \text{ en } p(\lambda).$$

EJERCICIOS (SECCIÓN 1, CAPÍTULO 6)

1. Sea A la matriz

$$A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

Determine cuáles de los siguientes vectores de $\mathbf{R}^2$ son vectores propios de A. En tal caso, determine el valor propio asociado.

a) (3, -1)
b) (2, 2)
c) (4, 4)
d) (3, -3)
e) (-10, 10)
f) (2, 1)

2. Sea A una matriz de orden n, suponga que para un vector no nulo $X \in \mathbf{R}^n$ se tiene $AX = 0$. Demuestre que $\lambda = 0$ es un valor propio de A.

3. Sea T: $\mathbf{R}^3 \to \mathbf{R}^3$ el operador lineal

$$T(x, y, z) = (x, 3x + 2y, 5x - 7y + 3z)$$

Determine cuáles de los siguientes vectores de $\mathbf{R}^3$ son vectores propios de T. En tal

caso, determine el valor propio asociado.

a) (3, 2, 0)	e) (4, 1, 3)	i) (0, 3, 21)
b) (0, 1, 7)	f) (−1, 3, 13)	j) (0, 0, 5)
c) (0, 0, 1)	g) (0, 0, 7)	k) (3, 4, 2)
d) (0, −2, −14)	h) (−2, 6, 26)	l) (1, 1, 1)

4. Sea A una matriz diagonal de orden n cuyos elementos de la diagonal principal son $\lambda_1, \lambda_2, \ldots, \lambda_n$. Determine el polinomio característico de A así como sus valores propios.
5. Sea A una matriz triangular de orden n. Determine el polinomio característico de A así como sus valores propios.
6. Sea V un espacio vectorial de dimensión finita. Determine el polinomio característico, los valores propios y los vectores propios del operador identidad en V, $Id: V \to V$ $Id(v) = v \ \forall\, v \in V$.
7. Sea V un espacio vectorial de dimensión finita. Determine el polinomio característico, los valores propios y los vectores propios del operador cero en V, $0: V \to V$, $0(v) = 0$, $\forall\, v \in V$.
8. ① Sea A una matriz de orden n, demuestre que A y A^t tienen el mismo polinomio característico y por lo tanto tienen los mismos valores propios. ¿Tienen los mismos vectores propios?
9. ① Pruebe que si λ es un valor propio de la matriz A, entonces λ^n es un valor propio de la matriz A^n, ($n \in \mathbb{N}$).
10. ① Sea A una matriz de orden n inversible, suponga que λ es un valor propio de A. En el ejercicio 18 se probará que $\lambda \neq 0$. Demuestre que λ^{-1} es un valor propio de A^{-1}. ¿Cuáles son los vectores propios de A^{-1}?
11. Sea A una matriz de orden n y sea c un número real. Si λ es un valor propio de A demuestre entonces que $c\lambda$ es un valor propio de cA.
12. Sea A una matriz de orden n, n impar. Compruebe que A posee al menos un valor propio (real).
13. Suponga que λ_1 es un valor propio de la matriz A y λ_2 es un valor propio de la matriz B. ¿Es $\lambda_1 + \lambda_2$ un valor propio de $A + B$?
14. Sea A una matriz cuadrada de orden n y c un número real. Si λ es un valor propio de A, demuestre que $\lambda + c$ es un valor propio de $A + cI$. Compare con el ejercicio anterior.
15. Sea A una matriz nilpotente de orden n (esto es, existe $k \in \mathbb{N}$ tal que $A^k = 0$). Compruebe que el polinomio característico de A es $p(\lambda) = (-1)^n \lambda^n$.
16. Sea A una matriz cuadrada de orden 2. Demuestre que el polinomio característico de A es:

$$p(\lambda) = \lambda^2 - (tr\,A)\lambda + det\,A$$

17. Sea A una matriz cuadrada de orden n. Pruebe que el término independiente del polinomio característico de A es $det\,A$.
18. Use el ejercicio anterior para demostrar que una matriz cuadrada de orden n es inversible si, y sólo si $\lambda = 0$ *no* es un valor propio de A. Obtenga una nueva demostración del ejercicio 2 a la luz de este ejercicio.
19. Sean A y B dos matrices de orden n.

a) Demuestre que si $I - AB$ es inversible, entonces $I - BA$ también lo es y

$$(I - BA)^{-1} = I + B(I - AB)^{-1}A$$

b) Use el inciso anterior para demostrar que las matrices AB y BA tienen los mismos valores propios.

20. Critique la siguiente "demostración" del teorema 1.3 (si A y B son matrices semejantes, entonces tienen el mismo polinomio característico): siendo A semejante a B existe una matriz inversible P tal que $A = P^{-1}BP$. Sea λ un valor propio de A, existe entonces un vector $v \neq 0$, tal que $Av = \lambda v$. Entonces, $P^{-1}BPv = \lambda v$, o sea $B(Pv) = \lambda(Pv)$. Como P es inversible, $Pv \neq 0$. Por lo tanto, λ es un valor propio de B (y Pv es un vector propio asociado a λ). Esto muestra que todo valor propio de A es también valor propio de B. Un argumento similar muestra que todo valor propio de B es también valor propio de A. Entonces A y B tienen los mismos valores propios. Sean $p_A(\lambda)$ y $p_B(\lambda)$ los polinomios característicos de A y B, respectivamente.

    Se ha demostrado entonces que $p_A(\lambda)$ y $p_B(\lambda)$ tienen las mismas raíces. Como ambos son polinomios del mismo grado, se concluye que $p_A(\lambda) = p_B(\lambda)$.
21. ① Sea B una matriz de orden n dada, considere el operador lineal $T: M_{n\times n} \to M_{n\times n}$, $T(A) = BA$. ¿Cómo son los valores propios de T respecto a los valores propios de B?
22. Sea A la matriz

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix}$$

a) Demuestre que el polinomio característico de A es:

$$p(\lambda) = -\lambda^3 + 3\lambda^2 + 11\lambda + 5$$

b) Compruebe que $p(A)$ es la matriz cero. Es decir, verifique que:

$$-A^3 + 3A^2 + 11A + 5I = 0$$

O sea que la matriz A "anula" a su polinomio característico. Éste es un hecho general, conocido como teorema de Hamilton-Cayley, que se demostrará en la sección 3 de este capítulo.

23. Para cada una de las siguientes matrices, determine
    1) su polinomio característico
    2) sus valores propios
    3) bases para los espacios propios correspondientes.

    Compruebe en cada caso que se satisface la conclusión del teorema 1.4.

a) $\begin{bmatrix} 5 & 0 \\ 0 & 6 \end{bmatrix}$ d) $\begin{bmatrix} 1 & 1 \\ -1 & 4 \end{bmatrix}$

b) $\begin{bmatrix} 5 & 0 \\ 7 & 6 \end{bmatrix}$ e) $\begin{bmatrix} -3 & 2 \\ 1 & -2 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ f) $\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 4 & 2 & 2 \end{bmatrix}$

g) $\begin{bmatrix} -3 & 2 & 1 \\ 0 & -5 & -7 \\ 0 & 2 & 4 \end{bmatrix}$ i) $\begin{bmatrix} 2 & 3 & 4 & 5 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 4 & 4 & 2 \end{bmatrix}$

h) $\begin{bmatrix} 2 & 8 & 7 \\ 0 & 3 & 1 \\ 0 & -1 & 1 \end{bmatrix}$ j) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 4 & 7 & 2 & 1 \\ 2 & 3 & 4 & 2 \end{bmatrix}$

24. Para cada uno de los siguientes operadores lineales determine

a) su polinomio característico

b) sus valores propios

c) bases para los espacios propios correspondientes.

Compruebe en cada caso que se satisface la conclusión del teorema 1.4 (para la matriz asociada al operador).

a) $T: P_1 \to P_1$, $T(a_0 + a_1x) = a_0 + (a_0 + a_1)x$.

b) $T: R^2 \to \mathbf{R}^2$, $T(x, y) = (x + y, y)$

c) $T: \mathbf{R}^3 \to \mathbf{R}^3$, $T(x, y, z) = (y, 2x + 4y - z, 3x + 5y - 2z)$

d) $T: M_{2\times 2} \to M_{2\times 2}$

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 4a + 5b & 2b \\ 2c + d & c + 3d \end{bmatrix}$$

e) $T: P_3 \to P_3$, $T(p) = p'$ (la derivada de p)

f) $M_{2\times 2} \to M_{2\times 2}$, $T(A) = A^t$

g) $T: P_3 \to P_3$, $T(a_0 + a_1x + a_2x^2 + a_3x_3)$

$= -2a_0 - 2a_1x + (8a_1 - 4a_2 + 2a_3)x^2 + (7a_1 + a_2 - 3a_3)x^3$

h) $T: M_{2\times 2} \to M_{2\times 2}$, $T(A) = BA$, en donde

$$B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

② 25. Sea V un espacio vectorial de dimensión 3 con el producto interno $(\cdot|\cdot)$. Sea $T: V \to V$ un operador lineal en V. Demuestre que existe una base ortonormal β de V respecto de la cual el operador T tiene una representación matricial de la forma

$$[T]_\beta = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & f & g \end{bmatrix}$$

(*Sugerencia*: demuestre primero que existe un vector propio $v \in V$ —véase ejercicio 12—. Complete luego a una base ortonormal de V.)

② 26. En este ejercicio se estudiará un cierto tipo de sistemas de ecuaciones diferenciales lineales con coeficientes constantes, los cuales son de la forma

$$\begin{aligned} x_1'(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \ldots + a_{1n}x_n(t) \\ x_2'(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \ldots + a_{2n}x_n(t) \\ &\vdots \\ x_n'(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \ldots + a_{nn}x_n(t) \end{aligned} \qquad (S)$$

en donde $x_1(t), x_2(t), \ldots, x_n(t)$ son ciertas funciones diferenciables de t definidas en algún intervalo I de $\mathbf{R}$ y a_{ij}, $1 \le i, j \le n$ son números reales dados. Al considerar el sistema (S), el objetivo es estudiar las soluciones de él, esto es, estudiar (de ser posible haciéndolas explícitas) las funciones $x_1(t), x_2(t), \ldots, x_n(t)$ que satisfacen el sistema (es decir, que sustituyéndolas en las ecuaciones del sistema, convierten a todas y cada una de éstas en identidades). Si se escribe:

$$X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

y se define la derivada de la matriz $X(t)$, denotada por $X'(t)$, como la matriz $n \times 1$ cuyos elementos son $x_i'(t)$, $i = 1, 2, \ldots, n$ (es decir, $X'(t)$ se obtiene derivando cada elemento de $X(t)$), se puede escribir el sistema (S) como

$$X'(t) = A\,X(t) \qquad (S')$$

Visto de esta manera el sistema (S), el objetivo es estudiar (describir) la matriz $X(t)$ que convierte a esta última expresión en una identidad. Para lograr esto, se procurarán soluciones del tipo $X(t) = e^{\lambda t}v$, en donde λ es algún número real y v es un vector de $\mathbf{R}^n$ (una matriz $n \times 1$). Sustituyendo en (S') se obtiene

$$\lambda e^{\lambda t}v = A e^{\lambda t}v$$

de donde $Av = \lambda v$ y se ve entonces que $X(t) = e^{\lambda t}v$ es una solución de (S') cuando λ es un valor propio de la matriz A y v es un vector propio asociado a λ. De la teoría de las ecuaciones diferenciales se sabe que si la matriz A tiene n valores propios (reales) distintos $\lambda_1, \lambda_2, \ldots, \lambda_n$, entonces la solución general del sistema (S') puede escribirse como

$$X(t) = c_1e^{\lambda_1 t}v_1 + c_2e^{\lambda_2 t}v_2 + \ldots + c_ne^{\lambda_n t}v_n$$

en donde $c_1, c_2, \ldots, c_n$ son constantes arbitrarias que pueden determinarse a partir de la condición inicial $X(0) = Xo$ (Xo es una matriz dada) y $v_1, v_2, \ldots, v_n$ son vectores propios asociados a los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectivamente.

Por ejemplo, estudie el sistema

$$\begin{aligned} x_1'(t) &= 5x_1(t) + 9x_2(t) \\ x_2'(t) &= 2x_1(t) + 2x_2(t) \end{aligned}$$

En este caso, la matriz $A = \begin{bmatrix} 5 & 9 \\ 2 & 2 \end{bmatrix}$ tiene los valores propios $\lambda_1 = 8$ y $\lambda_2 = -1$ y $v_1 = (3, 1)$ y $v_2 = (-3, 2)$ son vectores propios correspondientes a λ_1 y a λ_2, respectivamente (véase ejemplo 5). Entonces la solución general del sistema es:

$$X(t) = c_1e^{8t}\begin{bmatrix} 3 \\ 1 \end{bmatrix} + c_2e^{-t}\begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

(es decir, $x_1(t) = 3c_1e^{8t} - 3c_2e^{-t}$, $x_2(t) = c_1e^{8t} + 2c_2e^{-t}$). Si además se imponen las condiciones iniciales $x_1(0) = 0$, $x_2(0) = 3$ (o bien en términos de la matriz $X(t)$ la condición inicial $X(0) = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$), se obtiene:

$$X(0) = c_1\begin{bmatrix}3\\1\end{bmatrix} + c_2\begin{bmatrix}-3\\2\end{bmatrix} = \begin{bmatrix}0\\3\end{bmatrix}$$

de donde $c_1 = c_2 = 1$, y entonces se dice que la solución particular del sistema que satisface estas condiciones iniciales es:

$$X(t) = e^{8t}\begin{bmatrix}3\\1\end{bmatrix} + e^{-t}\begin{bmatrix}-3\\2\end{bmatrix}$$

a) Obtenga la solución general del sistema

$$x_1'(t) = x_1(t)$$
$$x_2'(t) = 2x_2(t)$$

b) Obtenga la solución del sistema

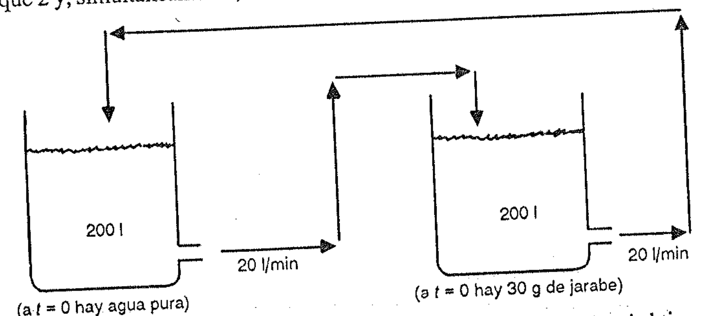
$$x_1'(t) = -2x_1(t) + 2x_2(t)$$
$$x_2'(t) = -15x_1(t) + 9x_2(t)$$

que satisfaga las condiciones iniciales $x_1(0) = x_2(0) = 0$.

c) Obtenga la solución general del sistema:

$$x_1'(t) = 9x_1(t) - 11x_2(t) - 5x_3(t)$$
$$x_2'(t) = 4x_1(t) - 4x_2(t) - 3x_3(t)$$
$$x_3'(t) = 2x_1(t) - 4x_2(t) + x_3(t)$$

③ 27. En el laboratorio de una compañía refresquera se tienen dos tanques llenos de 200 l de capacidad conectados entre sí. En el tanque 1 se tenía inicialmente agua pura y en el tanque 2 se tenían disueltos 30 g del concentrado de jarabe que se usa para la elaboración del refresco. Al tiempo $t = 0$ se abre un flujo de 20 l/min del tanque 1 al tanque 2 y, simultáneamente, otro flujo del tanque 2 al tanque 1 también de 20 l/min.



Llame $x_i(t)$ a la cantidad (en gramos) de jarabe presente en el tanque i al tiempo t, $i = 1, 2$. Observe que se tienen las siguientes condiciones iniciales: $x_1(0) = 0$ y $x_2(0) = 30$. Suponga que existe agitación ideal en el proceso, esto es, a medida que llega la solución a cada uno de los tanques, ésta queda inmediatamente incorporada en él.

a) Efectúe un "balance" de la cantidad de jarabe en cada uno de los tanques para concluir que las ecuaciones diferenciales que modelan el proceso son:

$$x_1'(t) = -0.1x_1(t) + 0.1x_2(t)$$
$$x_2'(t) = 0.1x_1(t) - 0.1x_2(t)$$

b) Resuelva el sistema anterior para obtener la solución general

$$X(t) = c_1\begin{bmatrix}1\\1\end{bmatrix} + c_2e^{-0.2t}\begin{bmatrix}1\\-1\end{bmatrix}$$

c) Imponga la condición inicial $X(0) = \begin{bmatrix}0\\30\end{bmatrix}$ para concluir que las funciones $x_1(t)$ y $x_2(t)$ que dan la cantidad de jarabe en los tanques 1 y 2, respectivamente al tiempo t son:

$$x_1(t) = 15 - 15e^{-0.2t}$$
$$x_2(t) = 15 + 15e^{-0.2t}$$

d) Observe que $\lim_{t\to\infty} x_1(t) = \lim_{t\to\infty} x_2(t) = 15$. Interprete este resultado.

e) Las normas de calidad de la compañía refresquera establecen que un refresco embotellado debe tener una concentración de 0.07 g de jarabe por cada litro de solución. Determine el tiempo en el que el tanque 1 llega a esta concentración. ¿El tanque 2 alcanza esta concentración en algún momento?

③ 28. La cinética química es la rama de la química que se encarga del estudio de las velocidades de una reacción. En este ejercicio se considerará un tipo de "reacción compleja" estudiada por la cinética química, cuya descripción matemática tiene que ver con lo estudiado en el ejercicio 26 anterior. Considere la reacción simple $A \to B$ en la que la sustancia A se descompone (reacciona) para producir la sustancia B. Se denotará por C_Y a la concentración molar de la sustancia Y; es decir, el número de moles presentes de Y por cada litro de solución. La llamada "ley de acción de masas" establece que la velocidad a la que se efectúa una reacción (la velocidad a la que A se consume o bien la velocidad a la que B se forma) es proporcional a la concentración molar de la(s) sustancia(s) reaccionante(s). En el caso de la reacción que se está considerando $A \to B$ esta ley establece que

$$C_A' = -kC_A$$

o bien,

$$C_B' = kC_B$$

en donde las derivadas se toman respecto del tiempo (es claro que C_A y C_B son funciones de t), y k es la constante de proporcionalidad (positiva) llamada "constante de velocidad de reacción", la cual es característica de cada reacción a condiciones dadas de presión y temperatura. Se escribe entonces la reacción como $A \overset{k}{\to} B$. En este ejercicio se estudiará la reacción

$$A \overset{k_1}{\to} R \overset{k_2}{\to} S$$

llamada "reacción consecutiva" en la cual la sustancia A se descompone para producir R y a su vez ésta sustancia se descompone para producir S. Suponga que a $t = 0$ se tenían No moles de A en un litro de solución reaccionante. Se tienen entonces tres funciones de t, a saber C_A, C_R y C_S y las condiciones iniciales $C_A(0) = No$, $C_R(0) = C_S(0) = 0$. El objetivo de este ejercicio es obtener expresiones explícitas para las funciones C_A, C_R y C_S que describen las concentraciones de cada uno de los reactivos y del producto final al tiempo t.

a) Use la ley de acción de masas para obtener las ecuaciones

$$C_A' = -k_1C_A$$
$$C_R' = k_1C_A - k_2C_R$$
$$C_S' = k_2C_R$$

Escriba este sistema de ecuaciones diferenciales como $X' = BX$ en donde

$$X = \begin{bmatrix} C_A \\ C_R \\ C_S \end{bmatrix} \qquad B = \begin{bmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & 0 \\ 0 & k_2 & 0 \end{bmatrix}$$

b) Demuestre que los valores propios de la matriz B son $\lambda_1 = 0$, $\lambda_2 = -k_1$, $\lambda_3 = -k_2$.

c) Demuestre que los vectores $v_1 = (0, 0, 1)$, $v_2 = (k_1 - k_2, -k_1, k_2)$ y $v_3 = (0, 1, -1)$ son vectores propios correspondientes a los valores propios $\lambda_1, \lambda_2, \lambda_3$, respectivamente.

d) Concluya de los dos incisos anteriores que la solución general del sistema planteado en el inciso a) es:

$$X = c_1\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + c_2e^{-k_1t}\begin{bmatrix} k_1 - k_2 \\ k_1 \\ k_2 \end{bmatrix} + c_3e^{-k_2t}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

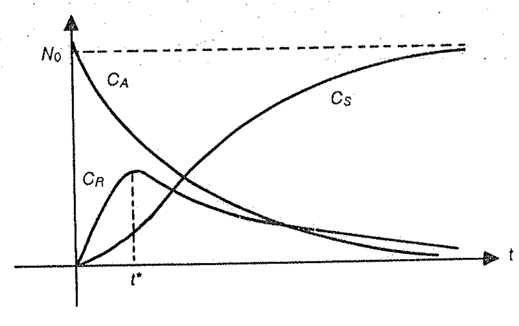
e) Imponga las condiciones iniciales $C_A(0) = No$, $C_R(0) = C_S(0) = 0$ para establecer los valores de las constantes

$$c_1 = No \qquad c_2 = \frac{No}{k_1 - k_2} \qquad c_3 = \frac{Nok_1}{k_1 - k_2}$$

f) Concluya entonces que las funciones que describen la concentración de las sustancias A, R y S son:

$$C_A = Noe^{-k_1t}$$
$$C_R = \frac{k_1}{k_2 - k_1}No(e^{-k_1t} - e^{-k_2t})$$
$$C_S = No - \frac{No}{k_2 - k_1}(k_2e^{-k_1t} - k_1e^{k_2t})$$

g) Demuestre que el comportamiento gráfico de las funciones establecidas en el inciso anterior es como sigue:



h) Demuestre que el punto t^* marcado en la gráfica anterior es:

$$t^* = \frac{ln(k_1/k_2)}{k_1 - k_2}$$

el cual corresponde a un máximo para C_R y a un punto de inflexión para C_S.

(**NOTA**: este ejercicio puede ser resuelto con una herramienta completamente distinta a la aquí presentada. Véase por ejemplo el libro del mismo autor, *Ecuaciones diferenciales, una introducción con aplicaciones*, Limusa, México, 1989, pp. 221-223.)

③ 29. Considere la reacción consecutiva

$$A \xrightarrow{k_1} R \xrightarrow{k_2} S \xrightarrow{k_3} T$$

en donde las constantes de velocidad de reacción valen $k_1 = 1$, $k_2 = 2$ y $k_3 = 3$. Sean C_A, C_R, C_S y C_T las concentraciones molares de las sustancias A, R, S y T, respectivamente al tiempo t. Suponga que a $t = 0$ se coloca 1 mol de A en un litro de solución reaccionante.

a) Demuestre que el sistema de ecuaciones diferenciales que describe el fenómeno es $X' = BX$ en donde:

$$X = \begin{bmatrix} C_A \\ C_R \\ C_S \\ C_T \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 0 & 2 & -3 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix}$$

b) Al resolver el sistema anterior, concluya que las funciones que describen las concentraciones molares de A, R, S y T al tiempo t son:

$$C_A = e^{-t}$$
$$C_R = e^{-t} - e^{-2t}$$
$$C_S = e^{-t} - 2e^{-2t} + e^{-3t}$$
$$C_T = 1 - 3e^{-t} + 3e^{-2t} - e^{-3t}$$

③ 30. Se llama reacción reversible (de primer orden) a una reacción del tipo

$$A \underset{k_2}{\overset{k_1}{\rightleftarrows}} R$$

en donde el reactante A produce R pero a su vez R produce A (las constantes de velocidad de reacción son k_1 y k_2, respectivamente). Sean C_A y C_R las concentraciones molares de las sustancias A y R, respectivamente. Suponga que a $t = 0$ se tenían Ao moles de A y Ro moles de R en un litro de solución reaccionante.

a) Use la ley de acción de masas para establecer las ecuaciones diferenciales que modelan el fenómeno

$$C_A' = -k_1C_A + k_2C_R$$
$$C_R' = k_1C_A - k_2C_R$$

Escriba este sistema como $X' = BX$ en donde:

$$X = \begin{bmatrix} C_A \\ C_R \end{bmatrix} \qquad B = \begin{bmatrix} -k_1 & k_2 \\ k_1 & -k_2 \end{bmatrix}$$

b) Demuestre que la matriz B tiene los valores propios $\lambda_1 = 0$, y $\lambda_2 = -(k_1 + k_2)$.

c) Demuestre que $v_1 = (k_2, k_1)$ y $v_2 = (1, -1)$ son vectores propios asociados a los valores propios λ_1 y λ_2, respectivamente.

d) Concluya que la solución general del sistema es:

$$X = c_1 \begin{bmatrix} k_2 \\ k_1 \end{bmatrix} + c_2 e^{-(k_1 + k_2)t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

e) Imponga las condiciones iniciales $C_A(0) = Ao$ y $C_R(0) = Ro$ para establecer los valores

$$c_1 = \frac{A_0 + R_0}{k_1 + k_2} \qquad c_2 = \frac{k_1 A_0 - k_2 R_0}{k_1 + k_2}$$

f) Concluya que las funciones C_A y C_R que describen las concentraciones de A y R al tiempo t son

$$C_A = \frac{A_0 + R_0}{k_1 + k_2} k_2 + \frac{k_1 A_0 - k_2 R_0}{k_1 + k_2} e^{-(k_1 + k_2)t}$$

$$C_R = \frac{A_0 + R_0}{k_1 + k_2} k_1 + \frac{k_1 A_0 - k_2 R_0}{k_1 - k_2} e^{-(k_1 + k_2)t}$$

g) Verifique que $\lim_{t \to \infty} C_A = \frac{A_0 + R_0}{k_1 + k_2} k_2$ y $\lim_{t \to \infty} C_R = \frac{A_0 + R_0}{k_1 + k_2} k_1$. A estos límites se les conoce como *concentraciones de equilibrio de las sustancias A y R*, respectivamente. Justifique esta nomenclatura e interprete físicamente estas concentraciones.

h) Sea x el número de moles por litro de A que han reaccionado al tiempo t. Demuestre que

$$x = \frac{k_1 A_0 - k_2 R_0}{k_1 + k_2} (1 - e^{-(k_1 + k_2)t})$$

i) Verifique que $\lim_{t \to \infty} x = \frac{k_1 A_0 - k_2 R_0}{k_1 + k_2}$. Dé un significado físico al valor de este límite.*

③ 31. Un caso de reacción compleja en la que se presentan las situaciones estudiadas en los ejercicios anteriores es una reacción consecutiva reversible del tipo:

$$A \underset{k_2}{\overset{k_1}{\rightleftarrows}} R \underset{k_4}{\overset{k_3}{\rightleftarrows}} S$$

Por simplicidad, suponga que $k_1 = k_2 = k_3 = k_4 = 1$. Sean C_A, C_R y C_S las concentraciones molares de las sustancias A, R y S al tiempo t. Suponga que al tiempo $t = 0$ se tenía 1 mol de A en un litro de solución reaccionante (en la que no se encontraban presentes R y S).

a) Use la ley de acción de masas para establecer las ecuaciones diferenciales

$$C_A' = -C_A + C_R$$
$$C_R' = C_A - 2C_R + C_S$$
$$C_S' = C_R - C_S$$

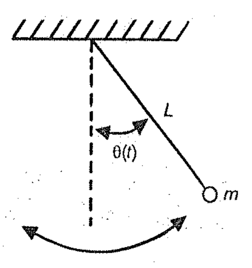
b) Resuelva el sistema para obtener:

$$C_A = \frac{1}{3} + \frac{1}{2} e^{-t} + \frac{1}{6} e^{-3t}$$

$$C_R = \frac{1}{3} - \frac{1}{3} e^{-3t}$$

$$C_S = \frac{1}{3} - \frac{1}{2} e^{-t} + \frac{1}{6} e^{-3t}$$

③ 32. En este ejercicio se estudiará el movimiento de un *péndulo simple*: un cuerpo de masa m sujeto a uno de los extremos de un hilo de longitud L con el otro extremo fijo



Sea $\theta = \theta(t)$ el ángulo que forma el hilo con la vertical al tiempo t. Suponga que a $t = 0$ se suelta la masa m de una posición para la cual $\theta = \theta o$. Se puede demostrar que si las oscilaciones de la masa m son pequeñas, la ecuación diferencial que describe el movimiento de ella es:

$$\frac{d^2\theta}{dt^2} + \frac{g}{L}\theta = 0 \qquad \text{(p)}$$

Observe además que se tienen las condiciones iniciales $\theta(0) = \theta o$ y $\theta'(0) = 0$ (esta última condición significa que la velocidad inicial de la masa m es nula). El objetivo es describir la función $\theta(t)$.

a) Introduzca las funciones $x_1(t)$ y $x_2(t)$ definidas como $x_1(t) = \theta(t)$, $x_2(t) = \theta'(t)$.

Demuestre que la ecuación (p) es equivalente al sistema de ecuaciones diferenciales

$$x_1'(t) = x_2(t)$$

$$x_2'(t) = -\frac{g}{L} x_1(t)$$

o bien, $X'(t) = AX(t)$, en donde

$$X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \qquad A = \begin{bmatrix} 0 & 1 \\ -g/L & 0 \end{bmatrix}$$

Observe que se tiene la condición inicial $X(0) = \begin{bmatrix} \theta_0 \\ 0 \end{bmatrix}$

b) Verifique que el polinomio característico de la matriz A es:

$$\lambda^2 + \frac{g}{L} = 0$$

y que entonces esta matriz tiene dos valores propios complejos

$$\lambda_1 = \sqrt{\frac{g}{L}}\, i \qquad \lambda_2 = -\sqrt{\frac{g}{L}}\, i$$

Se procederá formalmente usando el método descrito en el problema 26 para encontrar $X(t)$.

c) Demuestre que $v_1 = \left(1, \sqrt{\frac{g}{L}}\, i\right)$ y $v_2 = \left(-1, \sqrt{\frac{g}{L}}\, i\right)$ son vectores propios asociados a lo valores propios λ_1 y λ_2, respectivamente.

d) Concluya que la solución general del sistema es:

$$X(t) = c_1 \exp\left(\sqrt{\frac{g}{L}}\, it\right)\begin{bmatrix} 1 \\ \sqrt{\frac{g}{L}}\, i \end{bmatrix} + c_2 \exp\left(-\sqrt{\frac{g}{L}}\, it\right)\begin{bmatrix} -1 \\ \sqrt{\frac{g}{L}}\, i \end{bmatrix}$$

e) Imponga la condición inicial $X(0) = \begin{bmatrix} \theta_0 \\ 0 \end{bmatrix}$ para obtener los valores $c_1 = \frac{1}{2}\theta_0$, $c_2 - \frac{1}{2}\theta_0$.

f) Concluya entonces que la función $\theta(t)$ que describe el movimiento de la masa m es:

$$\theta(t) = \frac{1}{2}\theta_0 \exp\left(\sqrt{\frac{g}{L}}\, it\right) + \frac{1}{2}\theta_0 \exp\left(-\sqrt{\frac{g}{L}}\, it\right)$$

g) Use la fórmula de Euler $exp(\pm i\alpha) = \cos\alpha \pm i \operatorname{sen}\alpha$ para demostrar que $\theta(t)$ puede ser escrita como:

$$\theta(t) = \theta_o \cos\sqrt{\frac{g}{L}}\, t$$

Verifique que efectivamente esta función satisface la ecuación (p) junto con las condiciones iniciales dadas.

2. DIAGONALIZACIÓN

DEFINICIÓN 2.1 a) Se dice que la matriz cuadrada A es *diagonalizable* si existe una matriz inversible P tal que $P^{-1}AP$ es una matriz diagonal (es decir, si A es semejante a una matriz diagonal).

DEFINICIÓN 2.1 b) Se dice que el operador lineal $T: V \to V$ en el espacio vectorial V de dimensión finita es *diagonalizable* si existe una base β de V respecto de la cual la matriz que representa a T es una matriz diagonal.

Es claro que un operador lineal $T: V \to V$ es diagonalizable si, y sólo si la matriz que lo representa respecto de alguna (cualquiera) base de V lo es. En efecto, supóngase que T es diagonalizable y sea β' una base cualquiera de V. Sea $A = [T]_{\beta'}$. Según la definición 2.1 b) existe una base β de V tal que $B = [T]_\beta$ es una matriz diagonal. A y B son matrices semejantes y por tanto, existe una matriz inversible P tal que $B = P^{-1}AP$ (de hecho, P es la matriz de cambio de base de β a β'). Entonces, según la definición 2.1 a), A es diagonalizable. Recíprocamente, supóngase que A es diagonalizable y que A es la matriz que representa al operador $T: V \to V$ respecto de alguna base β' de V. Existe entonces una matriz P inversible tal que $P^{-1}AP$ es una matriz diagonal. Según el lema previo al teorema 4.1 (capítulo 4), existe una base β de V tal que P es la matriz de cambio de base de β a β'. Sea $B = [T]_\beta$. Como $P^{-1}AP = B$, B es una matriz diagonal, se concluye que T es diagonalizable.

COROLARIO

[De las definiciones 2.1 a) y 2.1. b).] Si la matriz A (el operador T) es diagonalizable, entonces el polinomio característico de A (de T) se puede factorizar como producto de factores lineales.

DEMOSTRACIÓN Si la matriz A es diagonalizable, A es semejante a una matriz diagonal D, dígase

$$D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

El polinomio característico de D es:

$$p(x) = det(D - xI) = det\begin{bmatrix} \lambda_1 - x & & & 0 \\ & \lambda_2 - x & & \\ & & \ddots & \\ 0 & & & \lambda_n - x \end{bmatrix}$$

$$= (\lambda_1 - x)(\lambda_2 - x)\dots(\lambda_n - x)$$

Siendo A semejante a D, el polinomio característico de A es también $p(x)$, el cual está factorizado como el producto de n (= orden de A) factores lineales.

Q.E.D.

En el siguiente teorema se establecen condiciones equivalentes al hecho de que una matriz A sea diagonalizable:

TEOREMA 2.1

> La matriz de orden n es diagonalizable si, y sólo si existe una base de $\mathbf{R}^n$ formada por vectores propios de A (equivalentemente: el operador T: $V \to V$, en donde V es un espacio vectorial de dimensión n, es diagonalizable si, y sólo si existe una base de V formada por vectores propios de T).

DEMOSTRACIÓN Supóngase primeramente que la matriz A es diagonalizable. Existe entonces una matriz inversible P tal que $P^{-1}AP = D$, en donde D es una matriz diagonal, dígase:

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

Claramente los valores propios de D son $\lambda_1, \lambda_2, \ldots, \lambda_n$, los cuales son también los valores propios de A, pues A es semejante a D.

Sean $v_j \in \mathbf{R}^n, j = 1, 2, \ldots, n$ los vectores columna de P. Siendo P una matriz inversible, estos vectores son linealmente independientes (véase teorema 4.3, capítulo 3).

En particular, todos ellos son vectores no nulos. Se verá que, de hecho, v_j es un vector propio de A asociado al valor propio λ_j, $j = 1, 2, \ldots, n$. Como $P^{-1}AP = D$, se tiene $AP = PD$. Igualando las columnas de estas dos matrices (AP y PD) se obtiene: j-ésima columna de $AP = Av_j = \lambda_j v_j =$ j-ésima columna de PD. Entonces, v_j es un vector propio de A asociado al valor propio λ_j. Se tiene entonces el conjunto de n vectores propios de A, $\beta = \{v_1, v_2, \ldots, v_n\}$ el cual se ha visto que es linealmente independiente. Como $dim\ \mathbf{R}^n = n$, estos vectores constituyen de hecho una base de $\mathbf{R}^n$.

Recíprocamente, supóngase que $\beta = \{v_1, v_2, \ldots, v_n\}$ es una base de $\mathbf{R}^n$ formada por vectores propios de A. Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los valores propios correspondientes. Sea P la matriz de orden n cuyos vectores columna son $v_1, v_2, \ldots, v_n$. Siendo $v_1, v_2, \ldots, v_n$ linealmente independientes, P es una matriz inversible. Como

$$Av_j = \lambda_j v_j \qquad j = 1, 2, \ldots, n$$

se tiene

$$AP = PD$$

en donde D es la matriz diagonal con elementos $\lambda_1, \lambda_2, \ldots, \lambda_n$ en su diagonal principal. Entonces, $P^{-1}AP = D$, mostrando así que A es diagonalizable.

En el caso de considerar al operador lineal T en lugar de la matriz A que lo representa, el argumento aquí presentado para la matriz A sufre modificaciones obvias, que se dejan a cargo del lector.

Q.E.D.

EJEMPLO 1 Investigue por ejemplo si la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & -2 \\ -1 & 1 & -3 \end{bmatrix}$$

es diagonalizable.

El polinomio característico de A es:

$$p(\lambda) = det(A - \lambda I) = det \begin{bmatrix} 1-\lambda & 0 & 0 \\ 1 & -\lambda & -2 \\ -1 & 1 & -3-\lambda \end{bmatrix} = (1-\lambda)\, det \begin{bmatrix} -\lambda & -2 \\ 1 & -3-\lambda \end{bmatrix}$$

$$= (1-\lambda)(\lambda^2 + 3\lambda + 2) = (1-\lambda)(1+\lambda)(2+\lambda)$$

Entonces, los valores propios de A son $\lambda_1 = 1$ $\lambda_2 = -1$ y $\lambda_3 = -2$.

El espacio propio E_{λ_1} es el espacio solución del sistema

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & -2 \\ -1 & 1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Éste es,

$$E_{\lambda_1} = \{(r, r, 0), r \in \mathbf{R}\}$$

Una base de E_{λ_1} es el vector propio $v_1 = (1, 1, 0)$.

El espacio propio E_{λ_2} es el espacio solución del sistema

$$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & -2 \\ -1 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

es decir,

$$E_{\lambda_2} = \{(0, 2r, r), r \in \mathbf{R}\}$$

Una base de este espacio es el vector propio $v_2 = (0, 2, 1)$.

Por último, el espacio propio E_{λ_3} es el espacio solución del sistema

$$\begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & -2 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Es decir,

$$E_{\lambda_3} = \{ (0, r, r), r \in \mathbf{R} \}$$

Una base de E_{λ_3} es el vector propio $v_3 = (0, 1, 1)$.

Se tienen entonces tres vectores propios $v_1 = (1, 1, 0)$, $v_2 = (0, 2, 1)$ y $v_3 = (0, 1, 1)$.

Si éstos fueran *linealmente independientes*, ellos constituirán una base de $\mathbf{R}^3$ y por lo tanto, A sería diagonalizable.

Obsérvese que la matriz P cuyos vectores columna son v_1, v_2 y v_3 es inversible pues

$$det\, P = det \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} = 1 \neq 0$$

lo cual implica la independencia lineal de v_1, v_2 y v_3.

Se concluye entonces que A es diagonalizable. De la demostración del teorema 2.1, se ve que $P^{-1}AP$ debe ser entonces una matriz diagonal (en cuya diagonal principal deben aparecer los valores propios de A). En efecto,

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & -2 \\ -1 & 1 & -3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

EJEMPLO 2

Por otro lado, en el último ejemplo de la sección anterior se considera la matriz

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 3 & 3 & 1 \end{bmatrix}$$

Sus valores propios son $\lambda_1 = 1$ y $\lambda_2 = 2$ y sus espacios propios correspondientes son:

$$E_{\lambda_1} = \{ (r, 0, 0, 0), r \in \mathbf{R} \}$$

$$E_{\lambda_2} = \left\{ \left(\frac{4}{3}r, \frac{1}{3}r, 0, r \right), r \in \mathbf{R} \right\}$$

Es claro entonces que en este caso se puede obtener a lo más, 2 vectores propios de A *linealmente independientes*, dígase $v_1 = (1, 0, 0, 0) \in E_{\lambda_1}$ y $v_2 = (4, 1, 0, 3) \in E_{\lambda_2}$ (pues $dim\ E_{\lambda_1} = dim\ E_{\lambda_2} = 1$). Estos vectores no pueden constituir una base de $\mathbf{R}^4$. Se concluye entonces que la matriz A no es diagonalizable.

Supóngase ahora que la matriz A de orden n tiene n valores propios distintos. Para investigar si A es diagonalizable se obtiene un vector propio correspondiente a cada valor propio de A, teniendo así un conjunto de n vectores propios de A en el espacio $\mathbf{R}^n$. Para concluir que estos vectores forman una base de $\mathbf{R}^n$ (y que por tanto, A es diagonalizable) se tiene que verificar la independencia lineal de ellos. Sin embargo, con el teorema siguiente se podrá ahorrar ese trabajo.

TEOREMA 2.2

Supóngase que la matriz cuadrada A de orden n tiene m valores propios distintos $\lambda_1, \lambda_2, \ldots, \lambda_m$ ($m \leq n$). Si v_i es un vector propio asociado al valor propio λ_i, $i = 1, 2, \ldots, m$ entonces el conjunto $\{v_1, v_2, \ldots, v_m\}$ es linealmente independiente.

DEMOSTRACIÓN

Por inducción sobre m = número de valores propios distinto de A. El resultado es obviamente válido para $m = 1$. Supóngase válido el resultado para $m = k$ ($k < n$) y pruébese para $m = k + 1$.

Considérese la combinación lineal

$$c_1v_1 + c_2v_2 + \ldots + c_kv_k + c_{k+1}v_{k+1} = 0 \qquad (2.1)$$

(se tendrá que probar que $c_1 = c_2 = \ldots = c_{k+1} = 0$). Al multiplicar la expresión anterior (por la izquierda) por la matriz A se obtiene

$$c_1(Av_1) + c_2(Av_2) + \ldots + c_k(Av_k) + c_{k+1}(Av_{k+1}) = A0 = 0$$

pero $Av_i = \lambda_i v_i$, $i = 1, 2, \ldots, k+1$, de modo que esta expresión se ve como

$$c_1\lambda_1v_1 + c_2\lambda_2v_2 + \ldots + c_k\lambda_kv_k + c_{k+1}\lambda_{k+1}v_{k+1} = 0 \qquad (2.2)$$

Al multiplicar la expresión (2.1) por λ_{k+1} y restar el resultado de la expresión (2.2) se obtiene

$$c_1(\lambda_1 - \lambda_{k+1})v_1 + c_2(\lambda_2 - \lambda_{k+1})v_2 + \ldots + c_k(\lambda_k - \lambda_{k+1})v_k = 0$$

Por la hipótesis de inducción, esta última expresión implica que

$$c_i(\lambda_i - \lambda_{k+1}) = 0 \quad i = 1, 2, \ldots, k$$

Pero por hipótesis del teorema $\lambda_i \neq \lambda_{k+1}$, $i = 1, 2, \ldots, k$, esto es, $\lambda_i - \lambda_{k-1} \neq 0$.

Entonces,

$$c_i(\lambda_i - \lambda_{k+1}) = 0 \Rightarrow c_i = 0 \quad i = 1, 2, \ldots, k$$

Por tanto, la expresión (2.1) queda convertida en:

$$c_{k+1}v_{k+1} = 0$$

Como $v_{k+1} \neq 0$ (pues v_{k+1} es un vector propio de A) se concluye finalmente que $c_{k+1} = 0$.

Q.E.D.

COROLARIO

Si la matriz A de orden n tiene n valores propios distintos, entonces A es diagonalizable.

DEMOSTRACIÓN Al tomar un vector propio v_i para cada valor propio λ_i, $i = 1, 2, \ldots, n$, se obtiene un conjunto linealmente independiente (teorema 2.2) de n vectores propios de $\mathbf{R}^n$. Éstos constituyen entonces una base de $\mathbf{R}^n$ formada por vectores propios de A. Entonces, A es diagonalizable (teorema 2.1).

Q.E.D.

Obsérvese que, según el corolario anterior, la condición de que una matriz de orden n tenga n valores propios distintos, es una condición suficiente para que esa matriz sea diagonalizable. Sin embargo, tal condición no es necesaria, como lo muestra el siguiente ejemplo:

EJEMPLO 3 Considere la matriz A de orden 3

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

El polinomio característico de A es:

$$p(\lambda) = det(A - \lambda I) = det\begin{bmatrix} 1-\lambda & 1 & 1 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{bmatrix} = (1-\lambda)(2-\lambda)^2$$

Entonces A posee sólo *dos* valores propios $\lambda_1 = 1$ y $\lambda_2 = 2$.

El espacio propio E_{λ_1} es:

$$E_{\lambda_1} = \{ (r, 0, 0), r \in \mathbf{R} \}$$

En él se toma el vector propio $v_1 = (1, 0, 0)$.

El espacio propio E_{λ_2} es:

$$E_{\lambda_2} = \{ (r+s, r, s), r, s \in \mathbf{R} \}$$

En él se pueden tomar los dos vectores linealmente independientes $v_2 = (1, 1, 0)$ y $v_3 = (1, 0, 1)$.

Se comprueba que los vectores propios $v_1 = (1, 0, 0)$, $v_2 = (1, 1, 0)$ y $v_3 = (1, 0, 1)$ constituyen una base $\mathbf{R}^3$. Entonces, A es diagonalizable. De hecho, si P es la matriz cuyos vectores columna son v_1, v_2 y v_3 se tiene (tal como lo asegura la demostración del teorema 2.1)

$$P^{-1}AP = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Se ha mostrado entonces una matriz de orden 3 con sólo 2 valores propios distintos que es diagonalizable. Si se ve un poco más de cerca este ejemplo, uno se da cuenta de que A pudo diagonalizarse gracias a que en el espacio propio E_{λ_2} se pudieron tomar dos vectores propios linealmente independientes (que junto con aquel del espacio propio E_{λ_1} formaron una base de $\mathbf{R}^3$ de vectores propios de A). El valor propio λ_2 aparece como raíz de multiplicidad 2 en el polinomio característico de A y la dimensión del espacio propio E_{λ_2} es, como se observa, 2.

Éste es un hecho general que se establece en el siguiente teorema:

TEOREMA 2.3

Sea A una matriz cuadrada de orden n. Sean $\lambda_1, \lambda_2, \ldots, \lambda_k$ los valores propios distintos de A y $E_{\lambda_1}, E_{\lambda_2}, \ldots, E_{\lambda_k}$ los espacios propios correspondientes. Las siguientes afirmaciones son equivalentes:

1) A es diagonalizable.
2) El polinomio característico de A es:

$$p(x) = (\lambda_1 - x)^{\alpha_1}(\lambda_2 - x)^{\alpha_2}\ldots(\lambda_k - x)^{\alpha_k}$$

y $dim\ E_{\lambda_i} = \alpha_i$, $i = 1, 2, \ldots, k$

3) $dim\ E_{\lambda_1} + dim\ E_{\lambda_2} + \ldots + dim\ E_{\lambda_k} = n$

DEMOSTRACIÓN Se demostrará que 1) ⇒ 2) ⇒ 3) ⇒ 1).

1) ⇒ 2). Ya se había observado que si A es una matriz diagonalizable entonces su polinomio característico se puede factorizar por completo como un producto de factores lineales (corolario de las definiciones 2.1 a) y 2.1 b)). Por otra parte, la dimensión del espacio propio E_{λ_i} es la dimensión del espacio solución del sistema homogéneo $(A - \lambda_i I)X = 0$ (véase teorema 7.4 capítulo 4). Sea $D = P^{-1}AP$ la forma diagonal de A (en la diagonal principal de D aparecen los valores propios de A). Las matrices $A - \lambda_i I$ y $D - \lambda_i I$ son semejantes y por tanto su rango coincide (igual al rango del operador $T - \lambda_i I$ que ambas representan).

El rango de $D - \lambda_i I$ es el número de líneas no nulas en $D - \lambda_i I$. Este número es $n - \alpha_i$ (¿por qué?). Entonces $dim\ E_{\lambda_i} = n -$ rango $(A - \lambda_i I) = n$-rango$(D - \lambda_i I) = n - (n - \alpha_i) = \alpha_i$, como se quería probar.

2) ⇒ 3). Se tiene $dim\ E_{\lambda_1} + dim\ E_{\lambda_2} + \ldots + dim\ E_{\lambda_k} = \alpha_1 + \alpha_2 + \ldots + \alpha_k =$ grado de $p(x) = n$

3) $\Rightarrow$ 1). Sea $\beta_i = \{v_{i_1}, v_{i_2}, \ldots, v_{i_{\alpha_i}}\}$ una base del espacio propio E_{λ_i} (los vectores $v_{i_1}, v_{i_2}, \ldots, v_{i_{\alpha_i}}$ son vectores propios de A asociados al valor propio λ_i).

Se propone que la unión β de las bases β_i, $i = 1, 2, \ldots, k$, es una base de $\mathbb{R}^n$.

Como por hipótesis $\alpha_1 + \alpha_2 + \ldots + \alpha_k = n = dim\ \mathbb{R}^n$, basta probar que los vectores de β son linealmente independientes. Tómese entonces la combinación lineal

$$\sum_{j=1}^{\alpha_1} c_{1_j} v_{1_j} + \sum_{j=1}^{\alpha_2} c_{2_j} v_{2_j} + \ldots + \sum_{j=1}^{\alpha_k} c_{k_j} v_{k_j} = 0 \qquad (2.3)$$

Llámese u_i, $i = 1, 2, \ldots, k$ al vector

$$u_i = \sum_{j=1}^{\alpha_i} c_{k_j} v_{k_j}$$

La expresión (2.3) se ve entonces como

$$u_1 + u_2 + \ldots + u_k = 0 \qquad (2.4)$$

Obsérvese que para $i = 1, 2, \ldots, k$ se tiene

$$Au_i = A\left(\sum_{j=1}^{\alpha_i} c_i v_{i_j}\right) = \sum_{j=1}^{\alpha_i} c_i A v_{i_j} = \sum_{j=1}^{\alpha_i} c_i \lambda_i v_{i_j} = \lambda_i \sum_{j=1}^{\alpha_i} c_i v_{i_j} = \lambda_i u_i \qquad (2.5)$$

(esto *NO* está diciendo que u_i sea un vector propio de A asociado al valor propio λ_i, pues para que esto fuese cierto se tendría que cumplir además que u_i sea no nulo. De hecho, se probará que $u_i = 0$, $i = 1, 2, \ldots, k$).

Multiplíquese la expresión (2.4) por la matriz A y úsese (2.5) para obtener

$$\lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_k u_k = 0 \qquad (2.6)$$

Al multiplicar (2.4) por λ_1 al restar el resultado a (2.6) se obtiene

$$(\lambda_2 - \lambda_1)u_2 + (\lambda_3 - \lambda_1)u_3 + \ldots + (\lambda_k - \lambda_1)u_k = 0 \qquad (2.7)$$

Multiplíquese esta última expresión por A y úsese (2.5) nuevamente para obtener

$$(\lambda_2 - \lambda_1)\lambda_2 u_2 + (\lambda_3 - \lambda_1)\lambda_3 u_3 + \ldots + (\lambda_k - \lambda_1)\lambda_k u_k = 0 \qquad (2.8)$$

Al multiplicar (2.7) por λ_2 y al restar el resultado a (2.8) se obtiene

$$(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)u_3 + \ldots + (\lambda_k - \lambda_1)(\lambda_k - \lambda_2)u_k = 0$$

Al continuar con este procedimiento, se llega a obtener un conjunto de expresiones

como las siguientes:

$$\begin{aligned} u_1 + u_2 + \ldots + u_k &= 0 \\ (\lambda_2 - \lambda_1)u_2 + \ldots + (\lambda_k - \lambda_1)u_k &= 0 \\ (\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)u_3 + \ldots + (\lambda_k - \lambda_1)(\lambda_k - \lambda_2)u_k &= 0 \\ \ldots \\ (\lambda_k - \lambda_1)(\lambda_k - \lambda_2)\ldots(\lambda_k - \lambda_{k-1})u_k &= 0 \end{aligned}$$

De la última expresión se obtiene que $u_k = 0$, pues por hipótesis se tiene que $(\lambda_k - \lambda_1)(\lambda_k - \lambda_2)\ldots(\lambda_k - \lambda_{k-1}) \neq 0$. Al sustituir en reversa en las expresiones anteriores, se obtiene que $u_{k-1} = u_{k-2} = \ldots u_2 = u_1 = 0$. Entonces, se tiene que:

$$u_i = \sum_{j=1}^{\alpha_i} c_{i_j} v_{i_j} = 0 \qquad i = 1, 2, \ldots, k$$

Como los vectores v_{i_j}, $j = 1, 2, \ldots, \alpha_i$ son linealmente independientes, esto implica que $c_{i_j} = 0$, $j = 1, 2, \ldots, \alpha_i$, $i = 1, 2, \ldots, k$. Esto muestra que los vectores de la base β son linealmente independientes. Por tanto, β es una base de $\mathbb{R}^n$ de vectores propios de A, lo que muestra que A es diagonalizable, como se quería.

Q.E.D.

Para finalizar esta sección, considérese el siguiente ejemplo:

EJEMPLO 4 Sea $T: P_3 \to P_3$ el operador lineal dado por

$$\begin{aligned} T(a_0 + a_1 + a_2x^2 + a_3x^3) = {} & 5a_0 + (2a_0 + 5a_1 + 3a_2)x + (-2a_0 + 2a_2)x^2 \\ & + (4a_0 + 6a_2 + 5a_3)x^3 \end{aligned}$$

Se pregunta si T es diagonalizable y, en tal caso, si lo es por una base de P_3 respecto de la cual la matriz que representa a T sea una matriz diagonal.

Tal como se observa al principio de esta sección, basta considerar el problema de diagonalización de una matriz que representa a T respecto de alguna base de P_3. Tome la base canónica de este espacio y obtenga la matriz que representa a T en esta base. Se tiene

$$\begin{aligned} T(1) &= 5 + 2x - 2x^2 + 4x^3 & (T(1))_\beta &= (5, 2, -2, 4) \\ T(x) &= 5x & (T(x))_\beta &= (0, 5, 0, 0) \\ T(x^2) &= 3x + 2x^2 + 6x^3 & (T(x^2))_\beta &= (0, 3, 2, 6) \\ T(x^3) &= 5x^3 & (T(x^3))_\beta &= (0, 0, 0, 5) \end{aligned}$$

Entonces la matriz $A = [T]_\beta$ es:

$$A = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 2 & 5 & 3 & 0 \\ -2 & 0 & 2 & 0 \\ 4 & 0 & 6 & 5 \end{bmatrix}$$

El polinomio característico de A (y por tanto, el de T) es:

$$p(\lambda) = det\,(A - \lambda I) = det \begin{bmatrix} 5-\lambda & 0 & 0 & 0 \\ 2 & 5-\lambda & 3 & 0 \\ -2 & 0 & 2-\lambda & 0 \\ 4 & 0 & 6 & 5-\lambda \end{bmatrix} = (5-\lambda)^3(2-\lambda)$$

Por tanto, T tiene sólo dos valores propios $\lambda_1 = 5$ y $\lambda_2 = 2$. Para determinar los espacios propios correspondientes, considérese el sistema homogéneo

$$\begin{bmatrix} 5-\lambda & 0 & 0 & 0 \\ 2 & 5-\lambda & 3 & 0 \\ -2 & 0 & 2-\lambda & 0 \\ 4 & 0 & 6 & 5-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Para $\lambda_1 = 5$ este sistema es:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 0 & 3 & 0 \\ -2 & 0 & -3 & 0 \\ 4 & 0 & 6 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = -\frac{3}{2}r, x_2 = s, x_3 = r, x_4 = t, r, s, t \in \mathbf{R}$. Entonces, el espacio propio E_{λ_1} es:

$$E_{\lambda_1} = \left\{ \left(-\frac{3}{2}r, s, r, t \right), \quad r, s, t \in \mathbf{R} \right\}$$

Obsérvese que

$$\left(-\frac{3}{2}r, s, r, t \right) = \frac{1}{2}r\,(-3, 0, 2, 0) + s(0, 1, 0, 0) + t(0, 0, 0, 1)$$

de modo que la dimensión del espacio E_{λ_1} es 3 [una base de él está constituida por los vectores propios (-3, 0, 2, 0), (0, 1, 0, 0), (0, 0, 0, 1)].

Para $\lambda_2 = 2$ el sistema homogéneo por resolver es:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 3 & 3 & 0 \\ -2 & 0 & 0 & 0 \\ 4 & 0 & 6 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = 0, x_2 = \frac{1}{2}r, x_3 = -\frac{1}{2}r, x_4 = r, r \in \mathbf{R}$. Entonces, el espacio propio E_{λ_2} es:

$$E_{\lambda_2} = \left\{ \left(0, \frac{1}{2}r, -\frac{1}{2}r, r \right), \quad r \in \mathbf{R} \right\}$$

La dimensión de E_{λ_2} es 1 [una base está constituida por el vector propio (0, 1, -1, 2)].

Obsérvese que

$$dim\ E_{\lambda_1} = 3 = \text{multiplicidad de } \lambda_1 \text{ en } p(\lambda)$$

$$dim\ E_{\lambda_2} = 1 = \text{multiplicidad de } \lambda_2 \text{ en } p(\lambda)$$

El teorema 2.3 asegura entonces que A (y por tanto, T) es diagonalizable.

Una base de $\mathbf{R}^4$ constituida por vectores propios de A es:

$$\{(-3, 0, 2, 0), (0, 1,0, 0), (0, 0, 0, 1), (0, 1, -1, 2)\}$$

En efecto, si P es la matriz que tiene a los vectores de esta base como vectores columna se tiene:

$$P^{-1}AP = \begin{bmatrix} -1/3 & 0 & 0 & 0 \\ 2/3 & 1 & 1 & 0 \\ 4/3 & 0 & 2 & 1 \\ -2/3 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 & 0 \\ 2 & 5 & 3 & 0 \\ -2 & 0 & 2 & 0 \\ 4 & 0 & 6 & 5 \end{bmatrix} \begin{bmatrix} -3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 2 & 0 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Al regresar al espacio P_3 se ve que

al vector (-3, 0, 2, 0) le corresponde el vector $-3 + 2x^2$

al vector (0, 1, 0, 0) le corresponde el vector x

al vector (0, 0, 0, 1) le corresponde el vector x^3

al vector (0, 1, -1, 2) le corresponde el vector $x - x^2 + 2x^3$

Por lo tanto, respecto de la base de P_3 dada por

$$\beta = \{-3 + 2x^2, x, x^3, x - x^2 + 2x^3\}$$

la representación matricial de T es:

$$[T]_\beta = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

EJERCICIOS (SECCIÓN 2, CAPÍTULO 6)

1. Demuestre que las siguientes matrices *no* son diagonalizables:

a) $\begin{bmatrix} 1 & -4 \\ 2 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 2 & 2 \\ -5 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 3 & 8 \\ -4 & 2 \end{bmatrix}$

d) $\begin{bmatrix} 1 & 8 & 7 \\ 0 & 1 & 4 \\ 0 & -5 & 2 \end{bmatrix}$

e) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 7 \\ 0 & 0 & -3 & 1 \end{bmatrix}$

[*Sugerencia*: use el corolario de las definiciones 2.1 a) y 2.1 b).]

2. Demuestre que la matriz

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

no es diagonalizable.

3. Demuestre que la matriz

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

es diagonalizable. Encuentre una matriz inversible P tal que $P^{-1}AP$ sea una matriz diagonal.

4. Demuestre que la matriz

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 1 & 4 \\ 3 & 4 & 2 \end{bmatrix}$$

es diagonalizable, mostrando que su polinomio característico posee tres raíces reales distintas.

5. Suponga que los polinomios de grado 5 mostrados a continuación son polinomios característicos de alguna matriz A de orden 5. En cada caso, diga si la matriz A *es* diagonalizable, si *no es* diagonalizable, o si *posiblemente sea* diagonalizable. En el último caso, diga qué condiciones adicionales se deben cumplir para que la matriz A sea diagonalizable.

a) $p(\lambda) = (1 - \lambda)(2 - \lambda)(3 - \lambda)(4 - \lambda)(5 - \lambda)$

b) $p(\lambda) = (1 - \lambda)^2(3 - \lambda)(-1 - \lambda)^2$

c) $p(\lambda) = -\lambda^5$

d) $p(\lambda) = (4 - \lambda)^3(\lambda^2 + \lambda + 1)$

e) $p(\lambda) = (-1 - \lambda)(-3 + 2\lambda + \lambda^2)(6 - 5\lambda + \lambda^2)$

f) $p(\lambda) = (1 - \lambda)^3(1 + \lambda^2)$

g) $p(\lambda) = (2 - \lambda)(\lambda^2 + 2\lambda + 6)(\lambda^2 - 3\lambda + 2)$

h) $p(\lambda) = (\lambda^3 - 5\lambda^2 + 6\lambda)(-20 + 9\lambda - \lambda^2)$

i) $p(\lambda) = (\lambda^4 - 1)(1 - \lambda)$

j) $p(\lambda) = -\lambda(\lambda^4 - 4\lambda^2 + 4)$

6. Sea A una matriz de orden n. Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los valores característicos de A. Suponga que A es diagonalizable. Demuestre entonces que

$$tr\,A = \lambda_1 + \lambda_2 + \ldots + \lambda_n$$
$$det\,A = \lambda_1\lambda_2 \ldots \lambda_n$$

7. Considere el operador derivación $D: P_n \to P_n$, $D(p) = p'$. Demuestre que D no es diagonalizable.

8. Sea $T: M_{2\times2} \to M_{2\times2}$ el operador $T(A) = A^t$. ¿Es T diagonalizable?

9. Demuestre que los siguientes operadores $T: \mathbf{R}^3 \to \mathbf{R}^3$ son diagonalizables.

En cada caso encuentre una base β de $\mathbf{R}^3$ respecto de la cual la matriz que representa a T es una matriz diagonal. Calcule también en cada caso el determinante del operador T.

a) $T(x, y, z) = (2x, 3x + 2y + z, 4x + y + 2z)$

b) $T(x, y, z) = (5x - 3y + 2z, 6x - 4y + 4z, 4x - 4y + 5z)$

c) $T(x, y, z) = (7x - 12y + 6z, 10x - 19y + 10z, 12x - 24y + 13z$

d) $T(x, y, z) = (x + y - 2z, -x + 2y + z, y - z)$

e) $T(x, y, z) = (7x - 2y - 4z, 3x - 2z, 6x - 2y - 3z)$

10. Pruebe que los siguientes operadores $T: M_{2\times2} \to M_{2\times2}$ son diagonalizables. En cada caso, encuentre una base β de $M_{2\times2}$ respecto de la cual la matriz que representa a T es una matriz diagonal. Calcule también en cada caso el determinante del operador T:

a) $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 2a+2b & -a+5b \\ 2a-c & 4a-8c+3d \end{bmatrix}$

b) $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 2a-2c+3d & b-c+2d \\ 3b-3c-6d & 2b-2c+4d \end{bmatrix}$

c) $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a+6b+12c & 4b+6c \\ -3b-5c & -3b-6c+d \end{bmatrix}$

11. Para cada una de las siguientes matrices, determine la expresión general de su n-ésima potencia.

a) $A = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}$

b) $a = \begin{bmatrix} 4 & 3 \\ -2 & -1 \end{bmatrix}$

c) $A = \begin{bmatrix} 3 & -4 \\ 1 & -2 \end{bmatrix}$

d) $A = \begin{bmatrix} 7 & -12 & -2 \\ 3 & -4 & 0 \\ -2 & 0 & -2 \end{bmatrix}$

e) $A = \begin{bmatrix} -2 & 8 & 6 \\ -4 & 10 & 6 \\ 4 & -8 & -4 \end{bmatrix}$

(*Sugerencia*: primeramente diagonalice la matriz A. Sea D la matriz diagonal semejante a A. Entonces, $A = P^{-1}DP$. Por tanto, $A^n = P^{-1}D^nP$. Véase demostración del teorema 3.3 de la siguiente sección.)

① 12. En este ejercicio se expone un método para obtener aproximaciones racionales de raíces cuadradas de números naturales. Sea k un número natural dado, del cual se quiere obtener una aproximación racional de su raíz cuadrada. Considere la matriz

$$A = \begin{bmatrix} 1 & k \\ 1 & 1 \end{bmatrix}$$

Defina el vector $Z_n = (x_n, y_n) \in \mathbf{R}^2$ como

$$Z_n = A^n\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Entonces, $\frac{x_n}{y_n}$ es una aproximación racional de $\sqrt{k}$.

En los incisos a) y b) se ve funcionar el método en dos casos concretos. Los incisos c), d), e), f), g) y h) dan los pasos para demostrar la validez del método.

a) Demuestre que las primeras 5 aproximaciones racionales de $\sqrt{2}$ son:

$$\frac{3}{2}, \frac{7}{5}, \frac{17}{12}, \frac{41}{29}, \frac{99}{70}$$

b) Obtenga las primeras 5 aproximaciones racionales de $\sqrt{3}$.

c) Pruebe que la matriz A tiene por valores propios a

$$\lambda_{1,2} = 1 \pm \sqrt{k}$$

Por tanto, A es diagonalizable.

d) Demuestre que una base del espacio propio asociado al valor propio $\lambda_1 = 1 + \sqrt{k}$ es $\{(\sqrt{k}, 1)\}$ y una base del espacio propio asociado al valor propio $\lambda_2, = 1 - \sqrt{k}$ es $\{(-\sqrt{k}, 1)\}$.

e) Compruebe que la matriz P que diagonaliza a A, así como su inversa están dadas por:

$$P = \begin{bmatrix} k & -\sqrt{k} \\ 1 & 1 \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} \frac{1}{2\sqrt{k}} & \frac{1}{2} \\ -\frac{1}{2\sqrt{k}} & \frac{1}{2} \end{bmatrix}$$

f) Demuestre que

$$A = \begin{bmatrix} 1 & k \\ 1 & 1 \end{bmatrix} = P\begin{bmatrix} 1+\sqrt{k} & 0 \\ 0 & 1-\sqrt{k} \end{bmatrix}P^{-1}$$

g) Compruebe que x_n y y_n (las coordenadas del vector z_n definido al principio del ejercicio) son:

$$x_n = \frac{1}{2}\left[(1+\sqrt{k})^{n+1} + (1-\sqrt{k})^{n+1}\right]$$

$$y_n = \frac{1}{2\sqrt{k}}\left[(1+\sqrt{k})^{n+1} - (1-\sqrt{k})^{n+1}\right]$$

h) Demuestre finalmente que

$$\lim_{n \to \infty} \frac{x_n}{y_n} = \sqrt{k}$$

13. Una *sucesión de Fibonacci* es una sucesión de números enteros $(a_0, a_1, a_2, \ldots)$ en la que $a_{n+2} = a_{n+1} + a_n$, $n = 0, 1, 2, \ldots$ con $a_0 = 0$ y $a_1 = 1$.

a) Determine los primeros diez términos de una sucesión de Fibonacci (a_n).

b) Sea A la matriz

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Defina el vector X_n de $\mathbf{R}^2$ como $X_n = A^n\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $n = 0, 1, 2, \ldots$ Demuestre que $X_n = (a_{n+1}, a_n)$, en donde (a_n) es una sucesión de Fibonacci.

c) Determine una fórmula general para la n-ésima potencia de la matriz A.

d) Concluya, de los dos incisos anteriores que el término a_n de una sucesión de Fibonacci (a_n) es:

$$a_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

(¡Observe que a_n es un número entero!)

e) Demuestre que

$$\lim \frac{a_{n+1}}{a_n} = \frac{1+\sqrt{5}}{2}$$

(la llamada por los griegos "razón áurea").

f) Repita el ejercicio con la sucesión de Fibonacci (a_n) en donde a_0 y a_1 son números arbitrarios.

② 14. Se dice que la matriz B es una raíz cuadrada de la matriz A, lo cual se escribe como $B = \sqrt{A}$, si $B^2 = A$.

a) Encuentre 4 raíces cuadradas de la matriz

$$A = \begin{bmatrix} 10 & 9 \\ 4 & 10 \end{bmatrix}$$

b) Encuentre una raíz cuadrada de la matriz

$$A = \begin{bmatrix} 4 & 10 & 10 \\ 0 & 11 & 7 \\ 0 & -2 & 2 \end{bmatrix}$$

¿Cuántas raíces cuadradas tiene esta matriz?

3. EL POLINOMIO MÍNIMO DE UNA MATRIZ Y EL TEOREMA DE HAMILTON-CAYLEY

Sea A una matriz del espacio vectorial $M_{2\times 2}$. Claramente las matrices A^2, A^3 y A^4 pertenecen también a este espacio. Considérese el conjunto $\{I, A, A^2, A^3, A^4\}$ en $M_{2\times 2}$. Éste es un conjunto de 5 elementos en el espacio vectorial $M_{2\times 2}$ el cual es de dimensión 4. Se sabe entonces que tal conjunto debe ser *linealmente dependiente*. Esto significa que existen constantes c_0, c_1, c_2, c_3 y c_4, no todas nulas tales que

$$c_0 I + c_1 A + c_2 A^2 + c_3 A^3 + c^4 A^4 = 0$$

Considérese el polinomio

$$p(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4$$

Éste es un polinomio no nulo de grado menor o igual que 4. Para una matriz cuadrada arbitraria M, defínase $p(M)$ de la manera natural

$$p(M) = c_0 I + c_1 M + c_2 M^2 + c_3 M^3 + c_4 M^4$$

$p(M)$ es entonces una matriz cuadrada del mismo orden que M.

Según el análisis inicial se tiene entonces que

$$p(A) = c_0 I + c_1 A + c_2 A^2 + c_3 A^3 + c_4 A^4 = 0$$

(el símbolo 0 que aparece en el lado izquierdo de esta igualdad se refiere a la matriz cero del espacio $M_{2\times 2}$). Se dice en este caso que A es una matriz que anula al polinomio $p(x)$.

En general, si A es una matriz cuadrada de orden n, deben existir n^2 constantes $c_1, c_2, \ldots, c_{n^2}$ no todas nulas tales que

$$c_0 I + c_1 A + c_2 A^2 + c_{n^2} A^{n^2} = 0$$

pues el conjunto $\{I, A, A^2, \ldots, A^{n^2}\}$ es un conjunto de $n^2 + 1$ elementos en un espacio vectorial de dimensión n^2 y por tanto, es *linealmente dependiente*.

Esto significa entonces que para el polinomio (no nulo)

$$p(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_{n^2} x^{n^2}$$

se tiene

$$p(A) = c_0 I + c_1 A + c_2 A^2 + \ldots + c_{n^2} A^{n^2} = 0$$

A es, pues, una matriz que anula al polinomio $p(x)$.

En esta sección será de interés estudiar polinomios que son anulados por matrices. Más concretamente, se estudiará el problema de, dada una matriz A de orden n, encontrar un polinomio de menor grado que es anulado por A. Tal polinomio existe, según el análisis anterior. Además, su grado no excede a n^2. Uno de los resultados importantes que se verá en esta sección (teorema 3.4) es que el grado de tal polinomio de hecho no excede a n.

DEFINICIÓN 3.1 Sea A una matriz cuadrada de orden n. Se dice que $m(x)$ es un *polinomio mínimo* de A si $m(x)$ es un polinomio mónico y es, de entre todos los polinomios que son anulados por A, un polinomio del menor grado posible.

Por ejemplo, considérese la matriz

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

El polinomio

$$p_1(x) = -2 + x - 2x^2 - x^3 + x^4$$

es un polinomio mónico que es anulado por A.

En efecto,

$$P_1(A) = -2I + A - 2A^2 - A^3 + A^4$$
$$= -2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} - 2\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} - \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix} + \begin{bmatrix} 16 & 0 \\ 0 & 16 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Por otra parte, el polinomio

$$p_2(x) = 4 - 4x + x^2$$

también es un polinomio mónico que es anulado por A, pues

$$P_2(A) = 4I - 4A + A^2 = 4\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 4\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

El grado de $p_2(x)$ es menor que el grado de $p_1(x)$. Entonces, $p_1(x)$ no puede ser el polinomio mínimo de A.

Aún más, el polinomio

$$m(x) = -2 + x$$

también es un polinomio mínimo que es anulado por A, pues

$$m(A) = -2I + A = -2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Como $m(x)$ es un polinomio de grado 1, cualquier polinomio de grado menor al de $m(x)$ sería un polinomio de grado cero —un polinomio constante— el cual obviamente no es anulado por A. Es claro, entonces, que $m(x)$ es un polinomio mínimo de A.

Después del siguiente teorema se podrá decir que $m(x)$ es *el* polinomio mínimo de A.

TEOREMA 3.1 El polinomio mínimo de una matriz es único.

DEMOSTRACIÓN Sean $m_1(x)$ y $m_2(x)$ dos polinomios mínimos de A. Según la definición 3.1, $m_1(x)$ y $m_2(x)$ son polinomios mónicos del mismo grado, dígase

$$m_1(x) = a_0 + a_1x + \ldots + a_{k-1}x^{k-1} + x^k$$
$$m_2(x) = b_0 + b_1x + \ldots + b_{k-1}x^{k-1} + x^k$$

Se quiere probar que $a_0 = b_0, a_1 = b_1, \ldots, a_{k-1} = b_{k-1}$ (o sea, que $m_1(x) = m_2(x)$).

Sea $p(x) = m_1(x) - m_2(x)$. Entonces,

$$p(x) = (a_0 - b_0) + (a_1 - b_1)x + \ldots + (a_{k-1} - b_{k-1})x^{k-1}$$

Supóngase, para obtener una contradicción, que existe j, $1 \leq j \leq k-1$ tal que $a_j - b_j \neq 0$. Tómese el menor índice j con esta propiedad. Entonces,

$$p(x) = (a_0 - b_0) + (a_1 - b_1)x + \ldots + (a_j - b_j)x^j$$

en donde $(a_j - b_j) \neq 0$.

Sea

$$\tilde{p}(x) = \frac{1}{a_j - b_j}p(x)$$

Se tiene que

1) $\tilde{p}(x)$ es un polinomio mónico
2) $\tilde{p}(x)$ es anulado por A, pues

$$\tilde{p}(A) = \frac{1}{a_j - b_j}p(A) = \frac{1}{a_j - b_j}(m_1(A) - m_2(A)) = \frac{1}{a_j - b_j}0 = 0$$

Como el grado del polinomio $\tilde{p}(x)$ es menor que k = grado de $m_1(x)$ = grado de $m_2(x)$, las propiedades 1 y 2 mostradas anteriormente contradicen que $m_1(x)$ (o $m_2(x)$) es un polinomio mínimo de A.

Entonces, $a_j - b_j = 0, j = 1, 2, \ldots, k - 1$. Es decir, que $p(x) = 0$ y por tanto, $m_1(x) = m_2(x)$, como se quería.

Q.E.D.

En el siguiente teorema se establece que de hecho el polinomio mínimo de una matriz A es la pieza con la que se construyen todos los polinomios que son anulados por A. Más precisamente, se demostrará que si $p(x)$ es un polinomio que es anulado por A, entonces $p(x)$ no es más que un múltiplo del polinomio mínimo de A.

TEOREMA 3.2 El polinomio mínimo de A, $m(x)$, divide a todo polinomio $p(x)$ que es anulado por A.

DEMOSTRACIÓN Sea $p(x)$ un polinomio tal que $p(A) = 0$, se debe mostrar que $p(x) = q(x)m(x)$ para algún polinomio $q(x)$. Al usar el algoritmo de la división para polinomios, se puede escribir

$$p(x) = q(x)m(x) + r(x)$$

en donde $q(x)$ es el cociente de la división de $p(x)$ entre $m(x)$ y $r(x)$ es el resto. Recuérdese que el polinomio $r(x)$ tiene la propiedad de que: 1) $r(x) = 0$, o 2) grado $r(x) <$ grado $m(x)$. Se afirma que en este caso, $r(x)$ debe cumplir la primera de estas propiedades.

En efecto, supóngase que se cumple la segunda propiedad para $r(x)$, esto es, grado $r(x) <$ grado $m(x)$. Como

$$r(x) = p(x) - q(x)m(x)$$

se tiene

$$r(A) = p(A) - q(A)m(A) = 0 - q(A)\,0 = 0$$

Entonces, $r(x)$ sería un polinomio que es anulado por A, el cual puede suponerse mónico, de grado menor al de $m(x)$. Esto contradice el hecho de que $m(x)$ es el polinomio mínimo de A.

Se tiene entonces que $r(x) = 0$ y por tanto, $p(x) = q(x)m(x)$, como se quería.

Q.E.D.

EJEMPLO 1 Por ejemplo, se vio que para la matriz

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

los polinomios

$$p_1(x) = -2 + x - 2x^2 - x^3 + x^4$$

y

$$p_2(x) = 4 - 4x + x^2$$

son anulados por A. Se vio también que $m(x) = -2 + x$ era el polinomio mínimo de A. Se tiene entonces (tal como lo asegura el teorema 3.2) que

$$p_1(x) = -2 + x - 2x^2 - x^3 + x^4 = (1 + x^2 + x^3)(-2 + x) = q_1(x)m(x)$$

$$p_2(x) = 4 - 4x + x^2 = (-2 - x)(-2 + x) = q_2(x)m(x)$$

Dadas las matrices semejantes A y B, se sabe que ambas poseen el mismo polinomio característico. Enseguida se probará que también poseen el mismo polinomio mínimo. Esto será una consecuencia del teorema siguiente:

TEOREMA 3.3 Si A y B son matrices semejantes dígase que $A = C^{-1}BC$, y $p(x)$ es cualquier polinomio, entonces

$$p(A) = C^{-1}p(B)C$$

DEMOSTRACIÓN Primeramente, pruébese que para cualquier $n \in \mathbb{N}$ se tiene la relación

$$A^n = C^{-1}B^nC$$

por inducción. Para $n = 1$, el resultado es obvio (es hipótesis del teorema). Supóngase entonces válido el resultado para $n = k$ y pruébese para $n = k + 1$.

Se tiene que

$$A^{k+1} = AA^k = A(C^{-1}B^kC) = (C^{-1}BC)(C^{-1}B^kC) = C^{-1}B^{k+1}C$$

Sea ahora el polinomio

$$p(x) = a_0 + a_1x + \dots + a_mx^m$$

Entonces,

$$\begin{aligned} p(A) &= a_0I + a_1A + \dots + a_mA^m = a_0C^{-1}C + a_1C^{-1}BC + \dots + a_mC^{-1}B^mC \\ &= C^{-1}(a_0I + a_1B + \dots + a_mB^m)C \\ &= C^{-1}p(B)C \end{aligned}$$

Q.E.D.

COROLARIO Matrices semejantes tienen el mismo polinomio mínimo.

DEMOSTRACIÓN Sean A y B dos matrices semejantes. Existe entonces C inversible tal que $A = C^{-1}BC$. Denótese por $m_A(x)$ y $m_B(x)$ a los polinomios mínimos de A y B, respectivamente. Entonces, por el teorema anterior

$$m_B(A) = C^{-1}m(B)C = C^{-1}0C = 0$$

Por tanto, por el teorema 3.2 existe $q(x)$ tal que

$$m_B(x) = q(x)m_A(x)$$

de donde grado $m_B(x) \geq$ grado $m_A(x)$.

Un argumento similar nos conduce a que grado $m_A(x) \geq$ grado $m_B(x)$. Entonces,

$$\text{grado } m_A(x) = \text{grado } m_B(x)$$

$m_B(x)$ es, pues, un polinomio mínimo, del mismo grado que $m_A(x)$, que es anulado por A. Por el teorema 3.1, $m_B(x)$ *debe ser* $m_A(x)$.

Q.E.D.

El siguiente teorema es uno de los resultados clásicos del álgebra lineal. En él se establece que el polinomio característico de una matriz cuadrada A es un polinomio que es anulado por A.

TEOREMA 3.4

(TEOREMA DE HAMILTON-CAYLEY.) Sea A una matriz cuadrada de orden n y sea $p(x)$ su polinomio característico. Entonces $p(A) = 0$

DEMOSTRACIÓN Escríbase el polinomio característico de A como

$$p(x) = det(A - xI) = (-1)^n(a_0 + a_1 + \ldots + a_{n-1}x^{n-1} + x^n)$$

Considérese la matriz $A - xI$ de orden n. En el teorema 5.2 del capítulo se estableció la fórmula

$$(A - xI)\, adj(A - xI) = det(A - x)I \tag{3.1}$$

Recuérdese que los elementos de la matriz adjunta de $A - xI$ son los cofactores de los elementos de la matriz $A - xI$. Estos cofactores se obtienen calculando determinantes de submatrices de $A - xI$ de orden $n - 1$. Es fácil convencerse entonces que cada uno de los elementos de la matriz $adj(A - xI)$ es un polinomio de grado a lo más $n - 1$ en x. Es decir,

$$adj(A - xI) = (p_{ij}(x))_{i,j,=1,\ldots,n}$$

en donde

$$p_{ij}(x) = c_{ij_0} + c_{ij_1}x + c_{ij_2}x^2 + \ldots + c_{ij_{n-1}}x^{n-1}$$

Llámese B_k a la matriz de orden n que tiene por elementos a c_{ij_k} $(k = 0, 1, \ldots, n - 1)$.

Es claro entonces, que

$$adj(A - xI) = B_0 + B_1x + B_2x^2 + \cdots + B_{n-1}x^{n-1}$$

Se tiene entonces de la expresión (3.1) que

$$(A - xI)(B_0 + B_1x + B_2x^2 + \ldots + B_{n-1}x^{n-1}) = (-1)^n(a_0 + a_1x + \ldots + x^n)I$$

Al realizar las operaciones indicadas en esta expresión e igualar los coeficientes de las potencias similares de x se obtiene

$$\begin{aligned} AB_0 &= (-1)^n a_0 I \\ -B_0 + AB_1 &= (-1)^n a_1 I \\ &\vdots \\ -B_{n-2} + AB_{n-1} &= (-1)^n a_{n-1} I \\ -B_{n-1} &= (-1)^n I \end{aligned}$$

Al premultiplicar la segunda de estas expresiones por A, la tercera por A^2, etc., (la

última por A^n) se obtiene

$$\begin{aligned} AB_0 &= (-1)^n a_0 I \\ -AB_0 + A^2B_1 &= (-1)^n a_1 A \\ &\vdots \\ -A^{n-1}B_{n-2} + A^nB_{n-1} &= (-1)^n a_{n-1} A^{n-1} \\ -A^nB_{n-1} &= (-1)^n A^n \end{aligned}$$

Obsérvese que al sumar todas estas expresiones, todos los términos que aparecen en el lado izquierdo de éstas, se cancelarán unos con otros. Entonces, se obtiene

$$(-1)^n(a_0I + a_1A + \ldots + a_{n-1}A^{n-1} + A^n) = 0$$

o sea que $p(A) = 0$.

Q.E.D.

COROLARIO

El polinomio mínimo de una matriz divide a su polinomio característico.

DEMOSTRACIÓN Obvio, del teorema anterior y del teorema 3.2.

Q.E.D.

EJEMPLO 2 Vea un ejemplo.

Sea A la matriz $A = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & -1 \end{bmatrix}$

El polinomio característico de A es:

$$p(x) = det(A - xI) = det\begin{bmatrix} -2-x & 0 & 0 \\ 0 & 2-x & 2 \\ 0 & 2 & -1-x \end{bmatrix}$$

$$= (-2-x)\,det\begin{bmatrix} 2-x & 2 \\ 2 & -1-x \end{bmatrix} = (-2-x)(x^2 - x - 6) = (3-x)(2+x)^2$$

o bien

$$p(x) = -x^3 - x^2 + 8x + 12$$

Tal como lo asegura el teorema de Hamilton-Cayley se tiene

$p(a) = -A^3 - A^2 + 8A + 12I$

$$= \begin{bmatrix} -8 & 0 & 0 \\ 0 & 20 & 14 \\ 0 & 14 & -1 \end{bmatrix} - \begin{bmatrix} 4 & 0 & 0 \\ 0 & 8 & 2 \\ 0 & 2 & 5 \end{bmatrix} + 8\begin{bmatrix} -2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & -1 \end{bmatrix} + 12\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Para determinar el polinomio mínimo $m(x)$ de A, uno se apoya en el corolario del teorema 3.4: como $m(x)$ debe dividir a $p(x)$, se tienen entonces las siguientes posibilidades [recordando que $m(x)$ es un polinomio mónico].

1) $m_1(x) = x - 3$
2) $m_2(x) = x + 2$
3) $m_3(x) = (x + 2)^2 = x^2 + 4x + 4$
4) $m_4(x) = (x + 2)(x - 3) = x^2 - x - 6$
5) $m_5(x) = (x + 2)^2(x - 3)$

Se tiene que

$$m_1(A) = A - 3I = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & -1 \end{bmatrix} - 3\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -5 & 0 & 0 \\ 0 & -1 & 2 \\ 0 & 2 & -4 \end{bmatrix} \neq 0$$

$$m_2(A) = A + 2I = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 2 \\ 0 & 2 & 1 \end{bmatrix} \neq 0$$

$$m_3(A) = A^2 + 4A + 4I = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 8 & 2 \\ 0 & 2 & 5 \end{bmatrix} - 4\begin{bmatrix} -2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & -1 \end{bmatrix} + 4\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 20 & 10 \\ 0 & 10 & 5 \end{bmatrix} \neq 0$$

$$m_4(A) = A^2 - A + 6I = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 8 & 2 \\ 0 & 2 & 5 \end{bmatrix} - \begin{bmatrix} -2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & -1 \end{bmatrix} + 6\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Entonces, $m_4(x) = x^2 - x + 6$ es el polinomio mónico de menor grado que es anulado por A. Éste es entonces el polinomio mínimo de A.

El siguiente teorema reduce considerablemente el trabajo de determinar el polinomio mínimo de una matriz a partir de su polinomio característico. Con él, se habría podido asegurar desde un principio que los polinomios $m_1(x)$, $m_2(x)$ y $m_3(x)$ del ejemplo anterior *no* podrían ser los polinomios mínimos de A.

La demostración del teorema se apoyará en el siguiente lema:

LEMA

Sea λ un valor propio de la matriz A, y sea X un vector propio asociado al valor propio λ. Si $q(x)$ es un polinomio cualquiera, se tiene que

$$q(A)X = q(\lambda)X$$

DEMOSTRACIÓN Primeramente se observa que para todo $k \in \mathbb{N}$ se tiene

$$A^k X = \lambda^k X$$

En efecto, para $k = 1$ esta relación establece que X es un vector propio de A asociado al valor propio λ (hipótesis del lema).

Supóngase válida la relación para k y pruébese para $k + 1$. Se tiene

$$A^{k+1}X = AA^kX = A\lambda^kX = \lambda^kAX = \lambda^k\lambda X = \lambda^{k+1}X$$

Sea ahora $q(x)$ el polinomio

$$q(x) = a_0 + a_1x + \ldots + a_mx^m$$

Entonces,

$$\begin{aligned} q(A)X &= (a_0I + a_1A + \ldots + a_mA^m)X \\ &= a_0X + a_1AX + \ldots + a_mA^mX \\ &= a_0X + a_1\lambda X + \ldots + a_m\lambda^mX \\ &= (a_0 + a_1\lambda + \ldots + a_m\lambda^m)X = q(\lambda)X. \end{aligned}$$

Q.E.D.

TEOREMA 3.5

Si $\lambda_1, \lambda_2, \ldots, \lambda_k$ son los valores propios distintos de la matriz A y si el polinomio característico de A es

$$p(x) = (-1)^n(x - \lambda_1)^{s_1} \ldots (x - \lambda_k)^{s_k} \qquad (3.2)$$

entonces el polinomio mínimo de A tiene la forma

$$m(x) = (x - \lambda_1)^{r_1} \ldots (x - \lambda_k)^{r_k} \qquad (3.3)$$

en donde $1 \leq r_i \leq s_i$, $i = 1, 2, \ldots, k$.

Es decir, todo factor lineal que aparezca en el polinomio característico de A, debe también aparecer en su polinomio mínimo.

DEMOSTRACIÓN Por el corolario del teorema 3.4 se sabe que si el polinomio característico de A es de la forma (3.2), entonces su polinomio mínimo debe ser de la forma (3.3) con $0 \le r_i \le s_i$, $i = 1, 2, \ldots, k$. Se debe probar entonces que $r_i \neq 0$ para todo $i = 1, 2, \ldots, k$.

Supóngase por contradicción, que $r_1 = 0$ (como los distintos factores de $m(x)$ conmutan, no hay pérdida de generalidad al suponer que es el primero de ellos el que tiene exponente cero). Entonces,

$$m(x) = (x - \lambda_2)^{r_2} \ldots (x - \lambda_k)^{r_k}$$

Sea X_1 un vector propio de A asociado al valor propio λ_1. Es decir,

$$AX_1 = \lambda_1 X_1$$

Según el lema previo a este teorema se tiene que

$$\begin{aligned} m(A)X_1 &= (A - \lambda_2 I)^{r_2} \ldots (A - \lambda_k I)^{r_k} X_1 \\ &= (\lambda_1 - \lambda_2)^{r_2} \ldots (\lambda_1 - \lambda_k)^{r_k} X_1 \end{aligned}$$

Obsérvese que el lado derecho de esta última expresión no es nulo, pues por una parte, $\lambda_1 \neq \lambda_j$, $j = 2, \ldots, k$ (por hipótesis) y por otra, el vector X_1 es un vector no nulo (pues es un vector propio de A). Entonces, se concluye que

$$m(A)X \neq 0$$

de donde $m(A) \neq 0$, hecho que contradice que $m(x)$ es el polinomio mínimo de A.

Q.E.D.

EJEMPLO 3 Considere por ejemplo la matriz

$$A = \begin{bmatrix} 2 & 2 & 0 & 0 \\ -1 & 5 & 0 & 0 \\ 2 & 0 & -1 & 0 \\ 4 & 0 & -8 & 3 \end{bmatrix}$$

El polinomio característico de A es:

$$p(x) = det(A - xI) = (x - 3)^2(x - 4)(x + 1)$$

Entonces, el polinomio mínimo de A puede ser, según el teorema anterior

1) $m(x) = (x - 3)(x - 4)(x + 1)$

o

2) $m(x) = (x - 3)^2(x - 4)(x + 1)$

Verifique la primera alternativa

$$\begin{aligned} m(A) &= (A - 3I)(A - 4I)(A + I) \\ &= \begin{bmatrix} -1 & 2 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 2 & 0 & -4 & 0 \\ 4 & 0 & -8 & 0 \end{bmatrix} \begin{bmatrix} -2 & 2 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & 0 & -5 & 0 \\ 4 & 0 & -8 & -1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 0 & 0 \\ -1 & 6 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 4 & 0 & -8 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Entonces, efectivamente $m(x) = (x - 3)(x - 4)(x + 1)$ es el polinomio mínimo de A.

El último resultado que se verá en esta sección relaciona la propiedad de diagonalización de una matriz A con la estructura de su polinomio mínimo.

TEOREMA 3.6

> Sean $\lambda_1, \lambda_2, \ldots, \lambda_k$ los valores propios distintos de la matriz A de orden n. Esta matriz es diagonalizable si, y sólo si su polinomio mínimo es:
>
> $$m(x) = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_k)$$

DEMOSTRACIÓN Supóngase primeramente que A es diagonalizable. Existe entonces una matriz inversible P tal que $P^{-1}AP = D$, en donde D es una matriz diagonal. Aún más, en la diagonal principal de D aparecen los valores propios de A (el valor propio λ_i aparece $dim\ E_{\lambda_i}$ veces en la diagonal principal de D —véase teorema 2.3—). Entonces, el polinomio característico de D (y por tanto, el de A) es:

$$p(x) = (-1)^n(x - \lambda_1)^{\alpha_1}(x - \lambda_2)^{\alpha_2} \ldots (x - \lambda_k)^{\alpha_k}$$

en donde $\alpha_i = dim\ E_{\lambda_i}$, $i = 1, 2, \ldots, k$ y $\alpha_1 + \alpha_2 + \ldots + \alpha_k = n$.

Según el teorema 3.5, en el polinomio mínimo de D deben aparecer todos los factores $(x - \lambda_1), (x - \lambda_2), \ldots, (x - \lambda_k)$. Es claro que el polinomio (mónico) de menor grado posible en el que aparecen todos estos factores es:

$$m(x) = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_k)$$

Se verá que de hecho éste es el polinomio mínimo de D.

Basta verificar entonces que $m(D) = 0$.

Obsérvese que

$$m(D) = (D - \lambda_1 I)(D - \lambda_2 I) \ldots (D - \lambda_k I)$$

A partir de la estructura de la matriz D, se puede deducir que $D - \lambda_1 I$ es una matriz diagonal con ceros en la primeras $dim\ E_{\lambda_1}$ posiciones de la diagonal principal, $D - \lambda_2 I$ es una matriz diagonal con ceros en las $dim\ E_{\lambda_2}$ posiciones después de las primeras $dim\ E_{\lambda_1}$ posiciones de la diagonal principal, etc. $D - \lambda_k I$ es una matriz diagonal con ceros en las últimas $dim\ E_{\lambda_k}$ posiciones de la diagonal principal.

Por otra parte, $q(D)$ es el producto de k matrices diagonales. Se sabe que al multiplicar matrices diagonales se obtiene también una matriz diagonal en cuya diagonal principal aparecen los productos de los elementos correspondientes de las diagonales principales de cada una de las matrices en el producto. Por lo analizado anteriormente acerca de la estructura de las matrices diagonales $D - \lambda_i I$, $i = 1, 2, \ldots, k$ que aparecen como vectores en $q(D)$, se concluye que $q(D) = 0$.

Entonces,

$$m(x) = (x - \lambda_i)(x - \lambda_2) \ldots (x - \lambda_k)$$

es el polinomio mínimo de D.

Es, por tanto, el polinomio mínimo de A (corolario del teorema 3.3).

Supóngase ahora que el polinomio mínimo de A es de la forma mostrada anteriormente para $m(x)$. Considérense los espacios propios $E_{\lambda_1}, E_{\lambda_2}, \ldots, E_{\lambda_k}$ correspondientes a los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_k$, respectivamente. Recuérdese que E_{λ_i} es el espacio solución del sistema homogéneo $(A - \lambda_i I)X = 0$. De acuerdo al teorema 7.6 del capítulo 4, la dimensión del espacio propio E_{λ_i} es:

$$dim\ E_{\lambda_i} = n - \text{rango}\,(A - \lambda_i I) \qquad i = 1, 2, \ldots, k$$

En el teorema 7.5 del capítulo 4 se probó que para cualesquiera dos matrices cuadradas A y B de orden n se cumple

$$\text{rango}\,(AB) \geq \text{rango}\,(B) + \text{rango}\,(A) - n$$

En el argumento que se presenta a continuación se hará uso repetido de las dos fórmulas anteriores.

Se tiene

$$\text{rango}\,(A - \lambda_1 I) = n - dim\ E_{\lambda_1}$$

$$\text{rango}\,(A - \lambda_2 I)(A - \lambda_1 I) \geq \text{rango}\,(A - \lambda_1 I) + \text{rango}\,(A - \lambda_2 I) - n$$
$$= n - dim\ E_{\lambda_1} - dim\ E_{\lambda_2}$$

$$\text{rango}\,(A - \lambda_3 I)(A - \lambda_2 I)(A - \lambda_1 I) \geq \text{rango}\,(A - \lambda_2 I)(A - \lambda_1 I) + \text{rango}\,(A - \lambda_3 I) - n$$
$$\geq n - dim\ E_{\lambda_1} - dim\ E_{\lambda_2} - dim\ E_{\lambda_3}$$

Al continuar con este argumento se llega a

$$\text{rango}\,(A - \lambda_k I) \ldots (A - \lambda_1 I) \geq n - (dim\ E_{\lambda_1} + \ldots + dim\ E_{\lambda_k})$$

Pero

$$(A - \lambda_k I) \ldots (A - \lambda_1 I) = m(A) = 0$$

y es claro que el rango de la matriz 0 es 0.

Se concluye entonces,

$$n \leq dim\ E_{\lambda_1} + \ldots + dim\ E_{\lambda_k}$$

Por otra parte, si α_i es la multiplicidad de λ_i como raíz en el polinomio característico de A, se tiene, de acuerdo al teorema 1.4

$$dim\ E_{\lambda_i} \leq \alpha_i \qquad i = 1, 2, \ldots, k$$

de donde

$$dim\ E_{\lambda_1} + \ldots + dim\ E_{\lambda_k} \leq \alpha_1 + \ldots + \alpha_k = n$$

En resumen se ha probado que

$$n = dim\ E_{\lambda_1} + \ldots + dim\ E_{\lambda_k}$$

y entonces, según el teorema 2.3, A es diagonalizable.

Q.E.D.

EJEMPLO 4 La matriz del ejemplo 3 previo al teorema 3.6

$$A = \begin{bmatrix} 2 & 2 & 0 & 0 \\ -1 & 5 & 0 & 0 \\ 2 & 0 & -1 & 0 \\ 4 & 0 & -8 & 3 \end{bmatrix}$$

tiene valores propios $\lambda_1 = 3$, $\lambda_2 = 4$ y $\lambda_3 = -1$. Se vio que su polinomio mínimo es:

$$m(x) = (x - 3)(x - 4)(x + 1)$$

Entonces, según el teorema 3.6, A es diagonalizable.

EJEMPLO 5 Por otro lado, la matriz

$$A = \begin{bmatrix} 1 & 4 & 0 & 2 & -4 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 6 & 1 & 0 & 0 \\ 0 & 3 & 2 & 1 & 0 \\ 0 & -1 & 1 & 3 & 1 \end{bmatrix}$$

tiene por polinomio característico a:

$$p(x) = det\,(A - xI) = det\begin{bmatrix} 1-x & 4 & 0 & 2 & -4 \\ 0 & 2-x & 0 & 0 & 0 \\ 0 & 6 & 1-x & 0 & 0 \\ 0 & 3 & 2 & 1-x & 0 \\ 0 & -1 & 1 & 3 & 1-x \end{bmatrix}$$

$$= (1-x)\,det\begin{bmatrix} 2-x & 0 & 0 & 0 \\ 6 & 1-x & 0 & 0 \\ 3 & 2 & 1-x & 0 \\ -1 & 1 & 3 & 1-x \end{bmatrix} = (1-x)^4(2-x)$$

Para que A sea diagonalizable, su polinomio mínimo tendría que ser:

$$m(x) = (x-1)(x-2)$$

pero

$$m(A) = (A-I)(A-2I) = \begin{bmatrix} 0 & 4 & 0 & 2 & -4 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 \\ 0 & 3 & 2 & 0 & 0 \\ 0 & -1 & 1 & 3 & 0 \end{bmatrix}\begin{bmatrix} -1 & 4 & 0 & 2 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 6 & -1 & 0 & 0 \\ 0 & 3 & 2 & -1 & 0 \\ 0 & -1 & 1 & 3 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 10 & 0 & -14 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 12 & -2 & 0 & 0 \\ 0 & 15 & 5 & -3 & 0 \end{bmatrix}$$

Se concluye entonces que la matriz A no es diagonalizable.

EJERCICIOS (SECCIÓN 3, CAPÍTULO 6)

1. Determine el polinomio mínimo de la matriz identidad de orden n así como el de la matriz cero de orden n.
2. Sea c un número real no nulo, considere la matriz de orden n, $A = cI$. Determine el polinomio característico de A así como su polinomio mínimo.
3. Sea A una matriz diagonal de orden n en cuya diagonal principal aparecen los elementos $\lambda_1, \lambda_2, \ldots, \lambda_n$. Demuestre que el polinomio mínimo de A es:

$$m(x) = (x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_n)$$

4. Sean α, β y γ tres números reales, considere la matriz

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \alpha & \beta & \gamma \end{bmatrix}$$

Demuestre que el polinomio característico de A es:

$$p(x) = -x^3 + \gamma x^2 + \beta x + \alpha$$

¿Cuál es el polinomio mínimo de A?

5. Para cada una de las siguientes matrices determine su polinomio característico y su polinomio mínimo.

a) $\begin{bmatrix} -5 & 0 \\ 8 & -6 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$

d) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}$

e) $\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$

f) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{bmatrix}$

g) $\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

h) $\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

i) $\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

6. Sea A una matriz de orden 3, demuestre que

$$m(x) = x^2 + 1$$

no puede ser el polinomio mínimo de A.

7. Sea A una matriz de orden n, suponga que A es nilpotente (existe $k \in \mathbb{N}$ tal que $A^k = 0$). Demuestre que $A^n = 0$. ¿Cuál es el polinomio mínimo de A?
(*Sugerencia*: véase el ejercicio 15 de la sección 1.)
8. Determine una matriz A de orden 2 cuyo polinomio mínimo sea $m(x) = x^2$.
9. Determine una matriz A de orden 3 cuyo polinomio mínimo sea $m(x) = x^2$.

② 10. Según el teorema 1.3 y el corolario del teorema 3.3, se tiene que si A y B son matrices semejantes, entonces ambas poseen el mismo polinomio característico y el mismo polinomio mínimo. Se puede demostrar que la afirmación recíproca *es falsa*, esto es, si A y B son dos matrices —del mismo orden— con el mismo polinomio característico y/o el mismo polinomio mínimo, esto no implica que A y B sean semejantes. Considere las matrices

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

a) Demuestre que A y B tienen el mismo polinomio característico.

b) Compruebe que A y B tienen el mismo polinomio mínimo.

c) Demuestre que A y B *no son semejantes*.

11. Sea D: $P_n \to P_n$ el operador lineal de derivación $D(p) = p'$. Determine el polinomio mínimo de la matriz que representa a D en alguna (en cualquiera, corolario del teorema 3.3) base de P_n.

12. Pruebe que las matrices dadas a continuación *no* son diagonalizables, demostrando que su polinomio mínimo no puede escribirse como un producto de factores lineales simples (teorema 3.6).

a) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

b) $\begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}$

c) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{bmatrix}$

d) $\begin{bmatrix} 4 & 6 & 6 \\ 1 & 3 & 2 \\ -1 & -5 & -2 \end{bmatrix}$

e) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

f) $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

g) $\begin{bmatrix} 2 & 0 & 1 & 3 \\ 4 & 2 & -1 & 1 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

h) $\begin{bmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$

13. Verifique la validez del teorema de Hamilton-Cayley para las matrices

a) $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$

b) $\begin{bmatrix} 3 & 1 & 3 \\ 1 & 0 & 0 \\ 3 & 1 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 2 & 3 & 1 & 4 \\ 2 & 1 & 1 & 1 \\ 3 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

14. Demuestre el siguiente resultado: Sea A una matriz $n \times n$. Sea

$$p(x) = (-1)^n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

el polinomio característico de A. Si $a_0 \neq 0$, entonces la matriz A es inversible y su inversa A^{-1} es:

$$A^{-1} = -\frac{1}{a_0}((-1)^n A^{n-1} + a_{n-1}A^{n-2} + \ldots + a_1 I)$$

(*Sugerencia*: véase el ejercicio 17 de la sección 1. Use el teorema de Hamilton-Cayley.)

15. Use el ejercicio anterior para determinar las inversas de las siguientes matrices:

a) $\begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 3 & 5 \\ 4 & 7 \end{bmatrix}$

c) $\begin{bmatrix} -1 & -1 \\ 5 & 3 \end{bmatrix}$

d) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

e) $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$

f) $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

16. Critique la siguiente "demostración" del teorema de Hamilton-Cayley:

Sea A una matriz cuadrada de orden n. El polinomio característico de A es $p(\lambda) = det(A - \lambda I)$.

Entonces, $p(A) = det(A - AI) = det(A - A) = det\, 0 = 0$.

4. DIAGONALIZACIÓN ORTOGONAL

Recuérdese que la matriz cuadrada A de orden n se dice ser *simétrica* si A coincide con su transpuesta A^t. Entonces, una matriz simétrica A es de la forma

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12} & a_{22} & \ldots & a_{2n} \\ & \vdots & & \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{bmatrix}$$

El objetivo de esta sección es mostrar que el problema de diagonalización de matrices simétricas siempre tiene solución. En otras palabras, se mostrará que toda matriz simétrica es semejante a una matriz diagonal.

DEFINICIÓN 4.1 Sea A una matriz cuadrada de orden n. Se dice que A es una matriz *diagonalizable ortogonalmente* si existe una matriz ortogonal P tal que $P^{-1}AP = (P^tAP)$ es una matriz diagonal.

Una matriz que es diagonalizable ortogonalmente es entonces una matriz diagonalizable [en el sentido de la definición 2.1 a)] para la cual la matriz inversible P (tal que $P^{-1}AP$ es una matriz diagonal) es una matriz ortogonal (es decir, $P^{-1} = P^t$).

Obsérvese que si la matriz A es diagonalizable ortogonalmente, entonces A es una matriz simétrica.

En efecto, si A es diagonalizable ortogonalmente existe una matriz ortogonal P tal que $P^tAP = D$, en donde D es una matriz diagonal (y por tanto simétrica). Entonces,

$$A = PDP^t$$

de donde

$$A^t = (PDP^t)^t = (P^t)^tD^tP^t = PDP^t = A$$

lo que muestra la afirmación.

El punto de interés que se presenta en esta sección es que la afirmación recíproca de la observación anterior es también verdadera. Es decir (se verá que) toda matriz simétrica es diagonalizable ortogonalmente (teorema 4.3).

El primer paso que se dará para lograr el objetivo es demostrar que una matriz simétrica tiene todos sus valores propios reales. Éste es el contenido del siguiente teorema.

TEOREMA 4.1

Sea A una matriz simétrica. Los valores propios de A son números reales.

DEMOSTRACIÓN Sea λ un valor propio cualquiera de A y sea X un vector propio asociado a λ. Se va a suponer que λ es un número complejo, dígase $\lambda = \alpha + i\beta$, $\alpha, \beta \in \mathbf{R}$ (se mostrará que $\beta = 0$). Obsérvese que en tal caso, el vector propio X, siendo solución —no trivial— del sistema homogéneo $(A - \lambda I)X = 0$ en el cual la matriz del sistema tiene elementos complejos, tiene en general coordenadas complejas. Se escribe entonces $X = (x_1, x_2, \ldots, x_n)$ en donde $x_1, x_2, \ldots, x_n \in C$.

Se tiene entonces

$$AX = \lambda X$$

Al tomar conjugación compleja en los elementos de las matrices involucradas en esta igualdad se obtiene

$$A\overline{X} = \overline{\lambda}\overline{X}$$

(se usa que $\overline{AX} = \overline{A}\,\overline{X} = A\overline{X}$ pues A es una matriz con elementos reales y $\overline{\lambda X} = \overline{\lambda}\overline{X}$.)

Al tomar transpuestas y al usar el hecho de que A es una matriz simétrica se obtiene

$$\overline{X}^tA = \overline{\lambda}\overline{X}^t$$

Al multiplicar por la izquierda por X y al usar $AX = \lambda X$ se obtiene

$$\overline{X}^t\lambda X = \overline{\lambda}\overline{X}^tX$$

de donde

$$(\lambda - \overline{\lambda})\overline{X}^tX = 0$$

Obsérvese que

$$\overline{X}^tX = [\overline{x}_1\,\overline{x}_2\ldots\overline{x}_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2$$

en donde $|x_i|$ denota el módulo del número complejo x_i, $i = 1, 2, \ldots, n$.

Entonces,

$$(\lambda - \overline{\lambda})(|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2) = 0$$

Como X es un vector propio de A, en particular es un vector no nulo. Esto permite deducir entonces de la expresión anterior que $\lambda - \overline{\lambda} = 0$. O sea,

$$(\alpha + i\beta) - (\alpha - i\beta) = 0$$

de donde $\beta = 0$, y por tanto $\lambda = \alpha \in \mathbf{R}$, como se quería.

Q.E.D.

La importancia del teorema 4.1 es que sitúa a todas las matrices simétricas como matrices "potencialmente diagonalizables" ya que en él se establece que el polinomio característico de la matriz simétrica A de orden n se puede siempre obtener (como raíces de él) n valores propios de A.

Se sabe que en general para una matriz A los vectores propios correspondientes a valores propios distintos son linealmente independientes (teorema 2.2). Un hecho interesante que se verá a continuación es que para una matriz simétrica, los vectores propios correspondientes a valores propios distintos no son sólo linealmente independientes sino que son además ortogonales (considerando el producto interno canónico de $\mathbf{R}^n$).

TEOREMA 4.2

Sean λ_i y λ_j dos valores propios distintos de la matriz simétrica A. Si X_i y X_j son vectores propios correspondientes a λ_i y λ_j respectivamente, entonces X_i es ortogonal a X_j [esto es, $(X_i \mid X_j) = 0$, en donde $(\cdot \mid \cdot)$ es el producto interno canónico de $\mathbf{R}^n$].

DEMOSTRACIÓN Se tiene entonces que

$$AX_i = \lambda_i X_i$$
$$AX_j = \lambda_j X_j$$

Al tomar transpuestas en la primera de estas expresiones y al usar el hecho de que A es simétrica se obtiene

$$X_i^t A = \lambda_i X_i^t$$

Al multiplicar por la izquierda por X_j y al usar la segunda de las expresiones iniciales se obtiene

$$X_i^t \lambda_j X_j = \lambda_i X_i^t X_j$$

de donde

$$(\lambda_j - \lambda_i)\, X_i^t X_j = 0$$

Obsérvese que

$$X_i^t X_j = [x_{1_i} x_{2_i} \dots x_{n_i}] \begin{bmatrix} x_{1_j} \\ x_{2_j} \\ \vdots \\ x_{n_j} \end{bmatrix} = x_{1_i} x_{1_j} + a_{2_i} x_{2_j} + \dots + x_{n_i} x_{n_j} = (X_i \mid X_j)$$

Se puede escribir entonces,

$$(\lambda_i - \lambda_j)(X_i \mid X_j) = 0$$

Como $\lambda_i \neq \lambda_j$ (por hipótesis) la expresión anterior obliga entonces a que $(X_i \mid X_j) = 0$, como se quería.

Q.E.D.

COROLARIO

Sea A una matriz simétrica de orden n. Si A tiene n valores propios distintos, entonces A es diagonalizable ortogonalmente.

DEMOSTRACIÓN Sean $\lambda_1, \lambda_2, \dots, \lambda_n$ los n valores propios distintos de A y sean $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n$ vectores propios correspondientes a cada uno de estos valores propios. Escríbase

$$X_i = \frac{\tilde{X}_i}{\|\tilde{X}_i\|} \qquad i = 1, 2, \dots, n$$

Es claro entonces, que X_i es un vector propio unitario correspondiente al valor propio λ_i, $i = 1, 2, \dots, n$.

Obsérvese que $\beta = \{X_1, X_2, \dots, X_n\}$ es una base ortonormal de $\mathbb{R}^n$ formada por vectores propios de A (la ortogonalidad del conjunto de vectores de β se deduce del teorema anterior). Sea β' la base canónica de $\mathbb{R}^n$ y sea P la matriz de cambio de base de β a β'. Entonces,

1) P tiene por vectores columna a los vectores propios $X_1, X_2, \dots, X_n$.
2) P es una matriz ortogonal (teorema 4.3 capítulo 5).

Entonces de la demostración del teorema 2.1 se concluye $P^{-1}AP = P^tAP = D$, en donde D es una matriz diagonal. Es decir, A es diagonalizable ortogonalmente.

Q.E.D.

EJEMPLO 1 Por ejemplo, considere la matriz simétrica

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix}$$

El polinomio característico de A es:

$$p(\lambda) = det(A - \lambda I) = det \begin{bmatrix} 1-\lambda & 0 & 2 \\ 0 & 2-\lambda & 0 \\ 2 & 0 & 4-\lambda \end{bmatrix} = \lambda(2-\lambda)(-5+\lambda)$$

Los valores propios de A son todos reales tal como lo asegura el teorema 4.1. Éstos son $\lambda_1 = 0$, $\lambda_2 = 2$ y $\lambda_3 = 5$.

Encuentre vectores propios correspondientes a cada uno de estos valores propios. Considere el sistema homogéneo.

$$\begin{bmatrix} 1-\lambda & 0 & 2 \\ 2 & 2-\lambda & 0 \\ 2 & 0 & 4-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Para $\lambda_1 = 0$ se tiene el sistema

$$\begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = -2r$, $x_2 = 0$, $x_3 = r$, $r \in \mathbb{R}$. Entonces, un vector propio asociado al valor propio $\lambda_1 = 0$ es $v_1 = (-2, 0, 1)$. Similarmente, halle que $v_2 = (0, 1, 0)$ es un vector propio asociado al valor propio $\lambda_2 = 2$ y $v_3 = (1, 0, 2)$ es un vector propio asociado al valor propio $\lambda_3 = 5$.

Tal como lo asegura el teorema 4.2 se tiene

$$(v_1 \mid v_2) = (-2)(0) + (0)(1) + (1)(0) = 0$$

$$(v_1 \mid v_3) = (-2)(1) + (0)(0) + (1)(2) = 0$$

$$(v_2 \mid v_3) = (0)(1) + (1)(0) + (0)(2) = 0$$

Sea $X_i \dfrac{1}{\| v_i \|} v_i$, $i = 1, 2, 3$. Entonces,

$$X_1 = \frac{1}{\|(-2, 0, 1)\|}(-2, 0, 1) = \left(-\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}\right)$$

$$X_2 = \frac{1}{\|(0, 1, 0\|}(0, 1, 0) = (0, 1, 0)$$

$$X_3 = \frac{1}{\|(1, 0, 2)\|}(1, 0, 2) = \left(\frac{1}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}\right)$$

La matriz P que tiene por vectores columna a X_1, X_2 y X_3 es:

$$P = \begin{bmatrix} -2/\sqrt{5} & 0 & 1/\sqrt{5} \\ 0 & 1 & 0 \\ 1/\sqrt{5} & 0 & 2/\sqrt{5} \end{bmatrix}$$

Tal como lo asegura la demostración del corolario del teorema 4.2 se tiene que P es una matriz ortogonal y que

$$P^tAP = \begin{bmatrix} -2/\sqrt{5} & 0 & 1/\sqrt{5} \\ 0 & 1 & 0 \\ 1/\sqrt{5} & 0 & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix} \begin{bmatrix} -2/\sqrt{5} & 0 & 1/\sqrt{5} \\ 0 & 1 & 0 \\ 1/\sqrt{5} & 0 & 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Es decir, A es diagonalizable ortogonalmente.

El siguiente (y último) objetivo en esta sección es eliminar del corolario del teorema 4.2 la hipótesis de que la matriz simétrica A tiene todos sus valores propios distintos. En el siguiente teorema se probará que cualquier matriz simétrica (con valores propios no necesariamente distintos) es diagonalizable ortogonalmente.

TEOREMA 4.3 Sea A una matriz simétrica de orden n. Entonces A es diagonalizable ortogonalmente.

DEMOSTRACIÓN La demostración es por inducción sobre n = orden de la matriz A.

Si $n = 1$, el teorema es obviamente verdadero, pues toda matriz de orden 1 es de hecho diagonal.

Supóngase entonces cierta la conclusión del teorema para toda matriz simétrica de orden $n - 1$ y pruébese su validez para la matriz A de orden n.

Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los valores propios de A (no se supone que estos valores sean distintos). Sea v_1 un vector propio asociado al valor propio λ_1. Existen vectores $v_2, v_3, \ldots, v_n \in \mathbf{R}^n$ tales que $\beta' = \{v_1, v_2, \ldots, v_n\}$ es una base de $\mathbf{R}^n$. Aplíquese el proceso de Gram-Schmidt a los vectores de la base β' para obtener la base ortonormal de $\mathbf{R}^n$

$$\beta_1 = \{X_1, X_2, \ldots, X_n\}$$

$X_1 = \dfrac{v_1}{\|v_1\|}$ es un vector propio unitario asociado al valor propio λ_1.

Sea B la matriz de cambio de base de β_1 a β_2, en donde β_2 es la base canónica de $\mathbf{R}^n$. Entonces B tiene por vectores columna a los vectores $X_1, X_2, \ldots, X_n$.

Además, B es una matriz ortogonal (véase teorema 4.3 capítulo 5).

Considérese el operador lineal T: $\mathbf{R}^n \to \mathbf{R}^n$ dado por $T(X) = AX$. La matriz de T respecto de la base canónica β_2 de $\mathbf{R}^n$ es A. Sea C la matriz de T respecto de la base β_1 de $\mathbf{R}^n$. Entonces,

$$C = B^{-1}AB = B^tAB$$

Obsérvese que

$$C^t = (B^tAB)^t = B^tA^t(B^t)^t = B^tAB = C$$

Entonces, C es una matriz simétrica.

Por otra parte, según la definición de la matriz C, la primera columna de esta matriz es:

$$(T(X_1))_{\beta_1} = (AX_1)_{\beta_1} = (\lambda_1 X_1)_{\beta_1} = (\lambda_1, 0, 0, \ldots, 0)$$

Como C es una matriz simétrica, ésta debe tener la siguiente forma:

$$C = \left[\begin{array}{c|ccc} \lambda_1 & 0 & 0 \;\ldots & 0 \\ \hline 0 & & & \\ 0 & & E & \\ \vdots & & & \\ 0 & & & \end{array}\right]$$

En donde E es una matriz simétrica de orden $n - 1$.

La matriz C es semejante a la matriz A y por tanto, ambas tienen el mismo polinomio característico. En particular, tienen los mismos valores propios. Esto implica entonces que la matriz simétrica E tiene por valores propios a $\lambda_2, \lambda_3, \ldots, \lambda_n$.

Según la hipótesis de inducción, E es una matriz diagonalizable ortogonalmente y por tanto, existe una matriz ortogonal F de orden $n - 1$, tal que $F^{-1}EF = F^tEF$ es una matriz diagonal. Más aún, la matriz diagonal F^tEF tiene los valores propios $\lambda_2, \lambda_3, \ldots, \lambda_n$ en su diagonal principal.

Sea G la matriz de orden n

$$G = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ 0 & & F & & \\ \vdots & & & & \\ 0 & & & & \end{bmatrix}$$

Si se usa la caracterización de matrices ortogonales: la matriz M es una matriz ortogonal si, y sólo si sus vectores columna forman una base ortonormal de $\mathbb{R}^n$ (véase ejercicio 16 de la sección 4 del capítulo 5), es fácil convencerse de que la matriz G es una matriz ortogonal (pues F lo es).

Como

$$G^{-1} = G^t = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ 0 & & F^t & & \\ \vdots & & & & \\ 0 & & & & \end{bmatrix}$$

Se ve que

$$G^{-1}CG = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ 0 & & F^t & & \\ \vdots & & & & \\ 0 & & & & \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ 0 & & E & & \\ \vdots & & & & \\ 0 & & & & \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ 0 & & F & & \\ \vdots & & & & \\ 0 & & & & \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ 0 & & F^tEF & & \\ \vdots & & & & \\ 0 & & & & \end{bmatrix} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

Sea

$$H = BG$$

Obsérvese que

$$H^{-1} = (BG)^{-1} = G^{-1}B^{-1} = G^tB^t = (BG)^t = H^t$$

Entonces H es una matriz (de orden n) ortogonal.

Finalmente,

$$H^{-1}AH = (BG)^{-1}A(BG) = G^{-1}(B^{-1}AB)G = G^{-1}CG = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

lo que muestra que A es diagonalizable ortogonalmente.

Q.E.D.

EJEMPLO 2 Vea con un ejemplo concreto cómo proceder para diagonalizar (ortogonalmente) una matriz simétrica que posee valores propios repetidos.

Sea A la matriz simétrica

$$A = \begin{bmatrix} 7 & -1 & -2 & 0 \\ -1 & 7 & 2 & 0 \\ -2 & 2 & 10 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

El polinomio característico de A es:

$$p(\lambda) = det(A - \lambda I) = det \begin{bmatrix} 7-\lambda & -1 & -2 & 0 \\ -1 & 7-\lambda & 2 & 0 \\ -2 & 2 & 10-\lambda & 0 \\ 0 & 0 & 0 & 6-\lambda \end{bmatrix}$$

$$= (6-\lambda)\, det \begin{bmatrix} 7-\lambda & -1 & -2 \\ -1 & 7-\lambda & 2 \\ -2 & 2 & 10-\lambda \end{bmatrix}$$

$$= (6-\lambda)(6-\lambda)^2(12-\lambda) = (6-\lambda)^3(12-\lambda)$$

Por tanto, A tiene solamente dos valores propios distintos $\lambda_1 = 6$ y $\lambda_2 = 12$.

Obtenga vectores propios correspondientes a estos valores propios. Para esto, considere el sistema homogéneo

$$\begin{bmatrix} 7-\lambda & -1 & -2 & 0 \\ -1 & 7-\lambda & 2 & 0 \\ -2 & 2 & 10-\lambda & 0 \\ 0 & 0 & 0 & 6-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Para el valor propio $\lambda_1 = 6$ obtenga el sistema

$$\begin{bmatrix} 1 & -1 & -2 & 0 \\ -1 & 1 & 2 & 0 \\ -2 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = r + 2s$, $x_2 = r$, $x_3 = s$, $x_4 = t$, $r, s, t \in \mathbb{R}$.

Del espacio propio E_{λ_1} (que tiene entonces dimensión 3) se pueden obtener los vectores propios linealmente independientes.

$$v_1 = (1, 1, 0, 0)$$
$$v_2 = (2, 0, 1, 0)$$
$$v_3 = (0, 0, 0, 1)$$

Al aplicar el proceso de Gram-Schmidt los vectores v_1, v_2, v_3 de la base del espacio E_{λ_1}, se obtienen los vectores

$$X_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)$$
$$X_2 = \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0\right)$$
$$X_3 = (0, 0, 0, 1)$$

que constituyen entonces una base ortonormal de E_{λ_1}.

Por otra parte, para el valor propio $\lambda_2 = 12$ resuelva el sistema

$$\begin{bmatrix} -5 & -1 & -2 & 0 \\ -1 & -5 & 2 & 0 \\ -2 & 2 & -2 & 0 \\ 0 & 0 & 0 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son $x_1 = -\frac{1}{2}r$, $x_2 = \frac{1}{2}r$, $x_3 = r$, $x_4 = 0$. Entonces, un vector propio asociado al valor propio $\lambda_2 = 12$ es $(-1, 1, 2, 0)$. Al normalizar este vector se obtiene

$$X_4 = \left(-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0\right)$$

Se tiene entonces, que la matriz P cuyos vectores columna son X_1, X_2, X_3 y X_4 es una matriz ortogonal (¿por qué?).

Finalmente se verifica que

$$p^tAP = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} & 0 \\ 0 & 0 & 0 & 1 \\ -1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} & 0 \end{bmatrix} \begin{bmatrix} 7 & -1 & -2 & 0 \\ -1 & 7 & 2 & 0 \\ -2 & 2 & 10 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & 0 & -1/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{3} & 0 & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 12 \end{bmatrix}$$

EJERCICIOS (SECCIÓN 4, CAPÍTULO 6)

① 1. Sea A una matriz simétrica de orden n. Suponga que todos los valores propios de A son cero. Muestre que en tal caso, A es la matriz cero. Por medio de un ejemplo concreto, muestre que esta afirmación es falsa si A no es simétrica.

① 2. Sea A una matriz simétrica de orden n, tal que en su diagonal principal aparecen sólo ceros. Demuestre que la suma de todos sus valores propios es igual a cero.

① 3. Compruebe que si dos matrices simétricas del mismo orden tienen el mismo polinomio característico, entonces son semejantes. Por medio de un ejemplo concreto, muestre que esta afirmación es falsa si las matrices no son simétricas.

4. Diagonalice ortogonalmente las siguientes matrices:

a) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

c) $\begin{bmatrix} -1 & 3 \\ 3 & -1 \end{bmatrix}$

d) $\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}$

e) $\begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}$

f) $\begin{bmatrix} 11 & 2 & -8 \\ 2 & 2 & 10 \\ -8 & 10 & 5 \end{bmatrix}$

g) $\begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$

h) $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{bmatrix}$

i) $\begin{bmatrix} 17 & -8 & 4 \\ -8 & 17 & -4 \\ 4 & -4 & 11 \end{bmatrix}$

j) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

5. Sea $T: \mathbb{R}^3 \to \mathbb{R}^3$ el operador lineal para el cual $T(e_1) = T(e_2) = T(e_3) = (1, 1, 1)$, en donde e_1, e_2, e_3 son los vectores de la base canónica de $\mathbb{R}^3$. Encuentre una base ortonormal de $\mathbb{R}^3$ respecto de la cual la representación matricial de T sea una matriz diagonal.

6. Sea $T: \mathbb{R}^4 \to \mathbb{R}^4$ el operador lineal para el cual $T(e_i) = e_{5-i}$, $i = 1, 2, 3, 4$, en donde $\{e_i, i = 1, 2, 3, 4\}$ es la base canónica de $\mathbb{R}^4$. Encuentre una base ortonormal de $\mathbb{R}^4$ respecto de la cual la representación matricial de T sea una matriz diagonal.

APÉNDICE SOBRE LA TEORÍA DE GRÁFICAS

La teoría de gráficas es una parte de la matemática en la que se estudian conjuntos de puntos y líneas y sus propiedades geométricas (no métricas). Los inicios de esta teoría se remontan al siglo XVIII con Leonard Euler (la figura más importante de la matemática de ese siglo). El primer problema que es reconocido propiamente como problema de la teoría de gráficas, conocido como el "problema de los siete puentes de Königsberg" fue propuesto (y resuelto) por Euler en un artículo enviado a la Academia de Ciencias de San Petersburgo el 26 de agosto de 1735, en el cual escribió:

> En la ciudad de Königsberg al este de Prusia, hay una isla llamada Kneiphoff, la cual está rodeada por una bifurcación de un río en la que hay 7 puentes [como se muestra en la figura]. La pregunta es si uno puede caminar a través de los 7 puentes cruzando todos y cada uno de ellos sólo una vez.

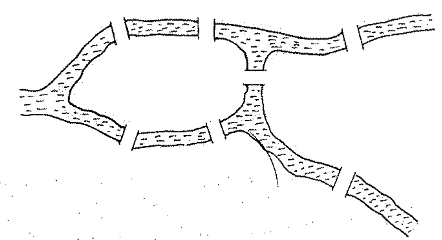


Figura 1

Desde el punto de vista de la teoría de gráficas, este problema es equivalente a preguntarse si la multigráfica* en la que cada línea representa un puente, puede ser recorrida completamente sin pasar dos veces por la misma línea y terminando el recorrido en el punto inicial del mismo. Euler demostró que tal recorrido no existe. Más aún, haciendo una generalización de esta situación demostró que para que una gráfica ("un conjunto de puntos y líneas que los unen") pueda ser recorrida pasando por todas y cada una de sus líneas y terminando en el punto inicial del recorrido,[†] es necesario y suficiente que la gráfica sea conexa (que esté constituida "por una sola pieza") y que cada uno de sus puntos sea incidente con un número

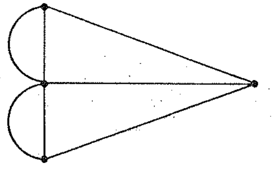


Figura 2

*Todos los términos técnicos usados en esta introducción serán definidos rigurosamente más adelante.

[†]Es decir, que puede trazarse sin levantar el lápiz del papel y terminando en el punto en donde se comenzó el trazo.

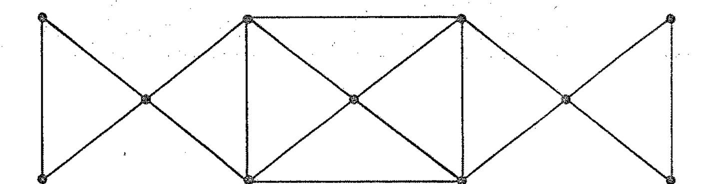


Figura 3

par de líneas (que cada punto de la gráfica esté conectado por un número par de líneas). A un recorrido en una gráfica como el que se ha mencionado se le conoce como "recorrido euleriano". Así entonces, la gráfica de la figura 3 tiene un recorrido euleriano mientras que la gráfica de la figura 4 no lo tiene. La multigráfica del problema de los siete puentes de Königsberg es conexa, pero tiene puntos (de hecho todos) incidentes con un número impar de líneas, por lo que ella posee un recorrido euleriano.

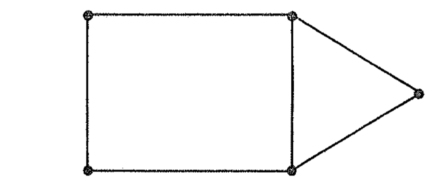


Figura 4

Algunas cuestiones de la teoría de gráficas pueden ser planteadas y resueltas con herramientas del álgebra lineal. En este apéndice se verán solamente algunos aspectos de esta conexión teoría de gráficas-álgebra lineal, en los cuales aparecen resultados que se piensa que además del valor intrínseco que ellos poseen son resultados sumamente bellos.

A.1. ALGUNAS DEFINICIONES PRELIMINARES

Una *gráfica* G está formada por un conjunto no vacío finito V de p *puntos* o *vértices* y un conjunto X de q pares no ordenados de puntos distintos de V, cada uno de los cuales es llamado *línea o lado* de G. Una gráfica tal es llamada una (p, q)-gráfica. De una manera obvia se asigna a cada (p, q)-gráfica G un diagrama en el que se muestren sus p puntos y sus q líneas. Tal diagrama es una *representación geométrica* de G. En realidad, se identifica siempre a la (p, q)-gráfica con alguna de sus representaciones geométricas.

EJEMPLO 1 El ejemplo más sencillo de gráfica es la $(1, 0)$-gráfica representada por



Figura 5

Esta gráfica es llamada gráfica trivial.

EJEMPLO 2 Como otro ejemplo vea la (5, 7)-gráfica G en la que

$$V = \{A, B, C, D, E\}$$

$$X = \{(A, B), (B, C), (C, D), (D, E), (E, A), (A, C), (C, E)\}$$

o bien, geométricamente

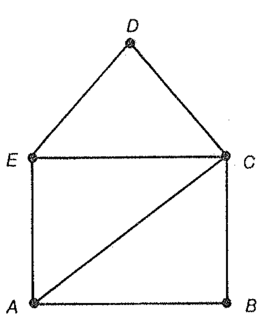


Figura 6

Observe que esta gráfica tiene muchas otras representaciones geométricas de aspecto distinto al de la figura 6, como por ejemplo

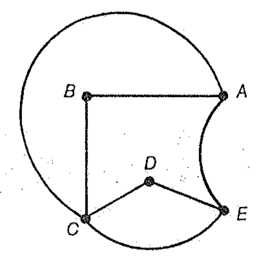


Figura 7

Se hace hincapié en que lo que constituye una gráfica G son los conjuntos V y X definidos anteriormente y no la posición relativa de sus vértices y forma de sus lados en alguna de sus representaciones geométricas. Así pues, los dos diagramas

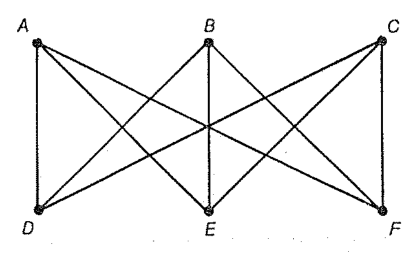


Figura 8

son representaciones geométricas distintas de la misma (6, 9)-gráfica G en la que

$$V = \{A, B, C, D, E, F\}$$

$$X = \{(A, D), (A, E), (A, F), (B, D), (B, E), (B, F), (C, D), (C, E), (C, F)\}$$

Puede ocurrir que en una gráfica dos de sus vértices estén conectados por varios lados, llamados *lados múltiples*, como por ejemplo en la figura 2. En tal caso, se dice que se trata de una *multigráfica*. Otro ejemplo de multigráfica es:

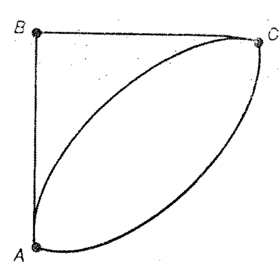


Figura 9

Observe también que en la definición de gráfica no se aceptan lados que unan un vértice consigo mismo, (llamados "loops"). Si se modifica la definición de gráfica de modo que se aceptan (lados múltiples y) "loops", se obtiene una *pseudográfica*, como por ejemplo:

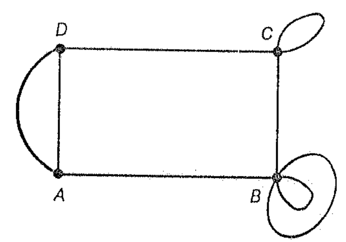


Figura 10

Si u y v son dos vértices de una gráfica G y (u, v) es un lado de G, se dice que u y v son *vértices adyacentes*. En tal caso, se dice que u y v son *vértices incidentes* en la línea (u, v) y que (u, v) es una *línea incidente* en los vértices u y v. Si dos líneas de G tienen un vértice común, se dice que son *líneas adyacentes*.

Por ejemplo, en la gráfica de la figura 6, se tiene que A y B son vértices adyacentes, mientras que A y D no lo son. Similarmente, las líneas (C, D) y (A, C) son adyacentes, mientras que las líneas (B, C) y (D, E) no lo son.

A.2. GRÁFICAS Y MATRICES

Dada una (p, q)-gráfica G, márquense sus vértices con índices $1, 2, \ldots, p$. Dígase entonces que los p vértices de G son $v_1, v_2, \ldots, v_p$. Se llamará G' a la gráfica marcada así obtenida a partir de G.

Una gráfica G obviamente tiene distintas maneras de marcarse. Por ejemplo, la (4, 5)-gráfica

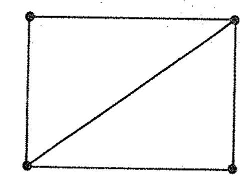
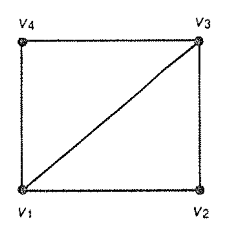
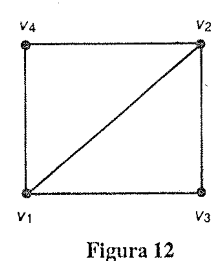
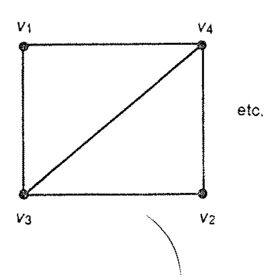


Figura 11

puede ser marcada como







etc.

Figura 12

Dada la (p, q)-gráfica G, márquense de algún modo los vértices de ella. Se asociará a la gráfica marcada G' así obtenida una matriz cuadrada A de orden p cuyos elementos a_{ij} están dados por

$$a_{ij} = \begin{cases} 1 & \text{si } v_i \text{ es adyacente a } v_j \\ 0 & \text{caso contrario} \end{cases}$$

A esta matriz se le llama *matriz de adyacencia* de la gráfica (marcada) G'. Una primera observación que se puede hacer de la matriz A es que ésta es una matriz *simétrica*, pues v_i es adyacente con v_j si, y sólo si v_j es adyacente con v_i.

EJEMPLO 3 Por ejemplo, considere la gráfica marcada

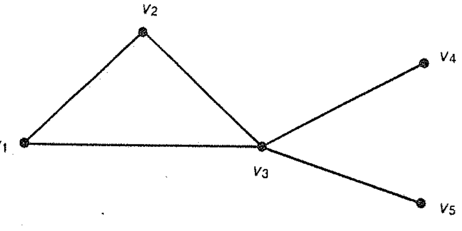


Figura 13

La matriz de adyacencia de esta gráfica es entonces

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

El interés final es asociar a cada gráfica G una matriz A de la cual se pueda obtener información acerca de G misma. Al proceder según el análisis anterior, se ve que existen distintas matrices asociadas a una misma gráfica G, pues la estructura de una matriz de adyacencia A depende de cómo es marcada la gráfica G. Por ejemplo, la (4, 5)-gráfica de la figura 11 tiene asociadas las matrices

$$A_1 = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

dependiendo si la gráfica es marcada como

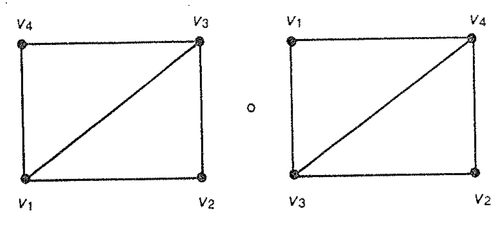


Figura 14

respectivamente.

Si se observan las gráficas marcadas de la figura 14 uno se dará cuenta de que la segunda puede obtenerse de la primera por medio de una permutación (en los índices) de sus vértices. Esta permutación es:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}$$

y así entonces el vértice v_i de la primera gráfica es el vértice $v_{\sigma(i)}$, $i = 1, 2, 3, 4$ en la segunda.

Considérese la matriz P de orden 4 asociada a la permutación σ*

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Obsérvese que la matriz P es una matriz ortogonal, de modo que $P^{-1} = P^t$. Además,

$$P^{-1}A_1P = P^tA_1P = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

*Recuérdese que dada una permutación σ del conjunto $\{1, 2, \ldots, n\}$, se asocia a ésta una matriz cuadrada de orden n, $P = (p_{ij})_{i, j = 1, \ldots, n}$ en donde

$$p_{ij} = \begin{cases} 1 & \text{si } j = \sigma(i) \\ 0 & \text{caso contrario} \end{cases}$$

Véase ejercicio 11 de la sección 1, capítulo 2.

$$= \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} = A_2$$

Resulta entonces que las matrices A_1 y A_2 asociadas a la misma (4, 5)-gráfica de la figura 11 son semejantes. Éste es un hecho general que se enuncia y se demuestra a continuación.

TEOREMA A1

> Sea G una (p, q)-gráfica, considere la gráfica marcada G' obtenida de G marcando sus vértices como $v_1, v_2, \ldots, v_p$. Sea σ una permutación del conjunto $\{1, 2, \ldots, p\}$. Si G'' es la gráfica marcada con los vértices $v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(p)}$ entonces,
>
> $$A_{G''} = P^t A_{G'} P$$
>
> en donde $A_{G'}$ y $A_{G''}$ son las matrices de adyacencia de las gráficas G' y G'', respectivamente y P es la matriz de la permutación σ.

DEMOSTRACIÓN Sea $A_{G'} = (a_{ij})_{i,\, j = 1, \ldots, p}$. Se sabe que los elementos p_{ij} de la matriz P son:

$$p_{ij} = \begin{cases} 1 & \text{si } j = \sigma(i) \\ 0 & \text{caso contrario} \end{cases}$$

Por lo tanto, la matriz P tiene en cada línea y en cada columna sólo un elemento no nulo, el cual es 1. Más precisamente, en la i-ésima línea todos los elementos son cero, excepto el de la $\sigma(i)$-ésima columna, el cual es 1.

Sea $A_{G'}P = (\theta_{ij})_{i,\, j = 1, \ldots, p}$. Entonces,

$$\theta_{ij} = \sum_{k=1}^{p} a_{ik} p_{kj} = a_{i\sigma^{-1}(j)}$$

La matriz $P^t A_{G'} P$ tiene en su i-ésima línea y j-ésima columna al elemento

$$\sum_{k=1}^{p} p_{ki} \theta_{kj} = \theta_{\sigma^{-1}(i)j} = a_{\sigma^{-1}(i)\sigma^{-1}(j)}$$

que es precisamente el correspondiente elemento de la matriz $A_{G''}$.

Q.E.D.

Se tiene entonces que todas las matrices de adyacencia asociadas a una misma gráfica G son semejantes. En particular esto permite hablar del "polinomio característico de una gráfica G", el cual puede ser definido como el polinomio característico de alguna matriz de adyacencia asociada a G. Sobre esto uno se ocupará más adelante.

A.3. CAMINOS

Sea G una gráfica. Un *camino* en G es una sucesión alternante de vértices y lados de G

$$v_0, (v_0, v_1), v_1, (v_1, v_2), v_2, \ldots, (v_{n-1}, v_n), v_n$$

que comienza y termina con vértices, y en la que cada lado es incidente con el vértice que le precede y con el que le sigue.

Al camino que comienza en v_0 y termina en v_n, pasando por $v_1, v_2, \ldots, v_{n-1}$, se le denota por $v_0v_1 \ldots v_n$ y se llama un $(v_0$-$v_n)$-camino. El $(v_0$-$v_n)$-camino se dice ser *cerrado* si $v_0 = v_n$ y *abierto* en caso contrario. Un camino en el que no hay repetición de vértices se llama *trayectoria*. Si en el camino no hay repetición de lados se llamará *paseo*. Un paseo cerrado en el que no hay repetición de vértices excepto el primero y el último se llama *ciclo*.

Por ejemplo, considérese la (8, 9)-gráfica marcada

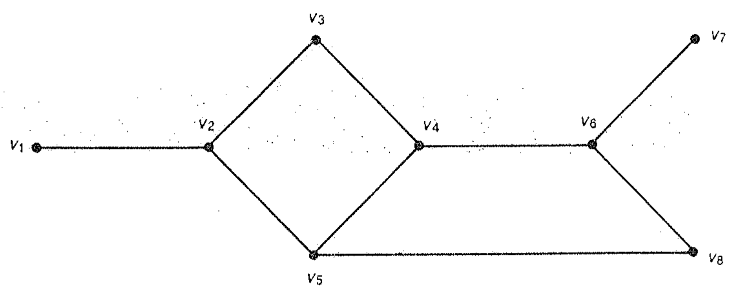


Figura 15

EJEMPLO 4 Los caminos $v_1v_2v_3v_4v_6v_7$ y $v_1v_2v_5v_8v_6v_7$ son ejemplos de $(v_1$-$v_7)$-caminos. Ambos son trayectorias. El camino $v_1v_2v_5v_4v_3v_2v_5v_8v_6v_7$ también es un $(v_1$-$v_7)$-camino pero no es trayectoria. El camino $v_1v_2v_3v_4v_5v_2$ es un $(v_1$-$v_2)$-camino el cual es un paseo pero no es trayectoria. Por último, los caminos $v_2v_3v_4v_5v_2$ y $v_5v_4v_6v_8v_5$ son ejemplos de ciclos.

Una gráfica G se dice ser *conexa* si cada par de puntos de ella pueden ser unidos por una trayectoria (si G está constituida por una sola pieza). Por ejemplo, la (5, 4)-gráfica

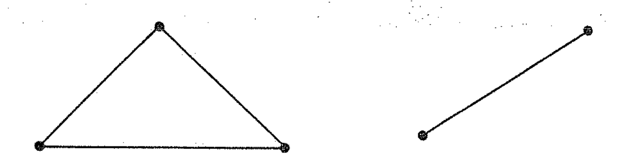


Figura 16

no es conexa.

La *longitud de un camino* es el número de lados que él contiene. Por ejemplo, respecto de la gráfica de la figura 15 se tiene que $v_1v_2v_3v_4$ y $v_1v_2v_5v_4$ son $(v_1\text{-}v_4)$-caminos de longitud 3. Similarmente, los caminos $v_2v_3v_4v_6v_7$ y $v_2v_5v_8v_6v_7$ son $(v_2\text{-}v_7)$-caminos de longitud 4.

Considérese una (p, q)-gráfica G con sus vértices marcados como $v_1, v_2, \ldots, v_p$. Dados dos vértices de G, dígase v_i y v_j, pregúntese por *cuántos* $(v_i\text{-}v_j)$-caminos existen, cuya longitud está previamente dada, dígase n. Se verá que es muy fácil dar respuesta a esta pregunta al usar la matriz de adyacencia de G. Pero antes véase un ejemplo.

Sea G la (5, 9)-gráfica

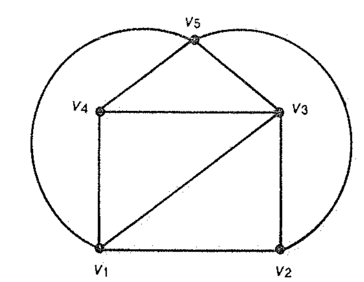


Figura 17

¿Cuántos caminos de longitud 2 unen a v_3 con v_5? No es difícil darse cuenta que solamente existen 3 de estos caminos, a saber $v_3v_4v_5$, $v_3v_2v_5$ y $v_3v_1v_5$.

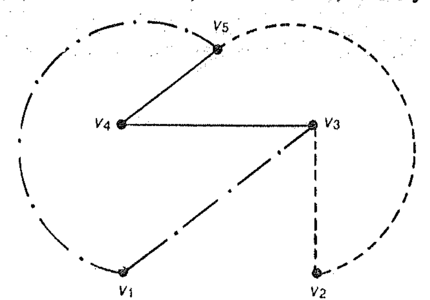


Figura 18

Sin embargo, investigar por ejemplo la cantidad de caminos de longitud 3 o 4 que unen a estos vértices v_3 y v_5, ya no es una labor tan sencilla. Se verá que existen 11 caminos de longitud 3 y 41 caminos de longitud 4.

TEOREMA A2

> Sea G una (p, q)-gráfica con sus vértices marcados $v_1, v_2, \ldots, v_p$. Sea A la matriz de adyacencia de G. Si M_{ij} es el elemento de la i-ésima línea y la j-ésima columna de la matriz A^n, existen M_{ij} caminos de longitud n que unen al vértice v_i con el vértice v_j en G.

DEMOSTRACIÓN Por inducción sobre n = longitud del camino que une al vértice v_i con el vértice v_j. Para $n = 1$ el resultado se sigue directamente de la manera como se construye la matriz de adyacencia de la gráfica G.

Supóngase entonces válido el resultado para $n = K$ y pruébese para $n = K + 1$.

Sea $A^K = (\alpha_{ij})_{i, j=1,\ldots,p}$. Entonces, el elemento que aparece en la i-ésima línea y j-ésima columna de la matriz $A^{K+1} = A^K A$ es:

$$\sum_{k=1}^{p} \alpha_{ik}a_{kj}$$

Se tiene que demostrar que ésta es la cantidad de $(v_i - v_j)$-caminos en G de longitud $K + 1$.

Sean $v_{k_1}, v_{k_2}, \ldots, v_{k_r}$ los vértices de G que son adyacentes con el vértice v_j.

Entonces, para ir del vértice v_i al vértice v_j por un camino de longitud $K + 1$, habrá que ir primeramente de v_i a algún v_{k_t}, $t = 1, 2, \ldots, r$ por un camino de longitud K, y posteriormente ir de v_{k_t} a v_j. Obsérvese además que

$$a_{kj} = \begin{cases} 1 \text{ si } k = k_1, k_2, \ldots, k_r \\ 0 \text{ caso contrario} \end{cases}$$

Entonces,

$$\sum_{k=1}^{p} \alpha_{ik}a_{kj} = \alpha_{ik_1} + \alpha_{ik_2} + \ldots + \alpha_{ik_r}$$

$\downarrow$ hipótesis de inducción

$$= \sum_{t=1}^{r} \text{(cantidad de caminos de longitud } K \text{ que unen a } v_i \text{ con } v_{k_t})$$

$$= \text{cantidad de caminos de longitud } K + 1 \text{ que unen a } v_i \text{ con } v_j.$$

Q.E.D.

EJEMPLO 5 Por ejemplo, la matriz de adyacencia de la (5, 9)-gráfica de la figura 17 es:

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Se tiene que

$$A^2 = \begin{bmatrix} 4 & 2 & 3 & 2 & 3 \\ 2 & 3 & 2 & 3 & 2 \\ 3 & 2 & 4 & 2 & 3 \\ 2 & 3 & 2 & 3 & 2 \\ 3 & 2 & 3 & 2 & 4 \end{bmatrix}.$$

y entonces, tal como se había visto directamente en G, existen 3 caminos de longitud 2 que unen a v_3 con v_5 (pues el elemento de la 3a. línea y 5a. columna de A^2 es 3). La información sobre la cantidad de cambios de longitud 3 y 4 que unen a v_i con v_j, $i, j = 1, \ldots, 5$, se obtiene de las matrices A^3 y A^4

$$A^3 = \begin{bmatrix} 10 & 10 & 11 & 10 & 11 \\ 10 & 6 & 10 & 6 & 10 \\ 11 & 10 & 10 & 10 & 11 \\ 10 & 6 & 10 & 6 & 10 \\ 11 & 10 & 11 & 10 & 10 \end{bmatrix} \qquad A^4 = \begin{bmatrix} 42 & 32 & 41 & 32 & 41 \\ 32 & 30 & 32 & 30 & 32 \\ 41 & 32 & 42 & 32 & 41 \\ 32 & 30 & 32 & 30 & 32 \\ 41 & 32 & 41 & 32 & 41 \end{bmatrix}$$

y así, pues, como ya se había dicho, existen 11 caminos de longitud 3 y 41 caminos de longitud 4 que unen a v_3 con v_5. ¿Podría describirlos explícitamente?

Establézcase ahora el concepto de "grado de un vértice": se llama *grado del vértice* v_i en una gráfica G, denotado por $grad\ v_i$, al número de líneas incidentes con v_i. Así por ejemplo, en la (5, 9)-gráfica de la figura 17 se tiene que

$$grad\ v_1 = grad\ v_3 = grad\ v_5 = 4 \qquad grad\ v_2 = grad\ v_4 = 3$$

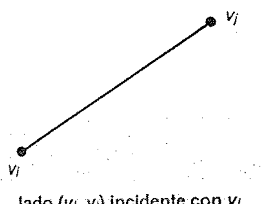
Con la ayuda de la matriz de adyacencia es posible obtener información acerca del grado de un vértice de una gráfica G. Así lo muestra el siguiente corolario del teorema A2:

COROLARIO

Sea G una (p, q)-gráfica con sus vértices marcados como $v_1, v_2, \ldots, v_p$. Sea A la matriz de adyacencia de G. Si $A^2 = (b_{ij})_{i,j=1,\ldots,p}$, entonces

$$grad\ v_i = b_{ii} \qquad i = 1, 2, \ldots, p$$

DEMOSTRACIÓN Según el teorema A_2, b_{ii} es la cantidad de caminos de longitud 2 que unen al vértice v_i con él mismo. Basta observar entonces que por cada lado incidente con v_i existe un (y sólo un) camino de longitud 2 que une a v_i con v_i



lado (v_i, v_j) incidente con v_i

v_j

v_i

$v_iv_jv_i$ camino de long. 2 que une a v_i con v_i

Figura 19

Q.E.D.

El siguiente resultado relaciona los valores de las trazas de las matrices A^2 y A^3 (A es la matriz de adyacencia de G) con la estructura de la gráfica G.

TEOREMA A3

Sea G una (p, q)-gráfica con sus vértices marcados como $v_1, v_2, \ldots, v_p$. Sea A la matriz de adyacencia de G. Entonces,

1) $tr\ A^2 = 2q$
2) $\frac{1}{6} tr\ A^3 =$ cantidad de ciclos de longitud 3 en G.

DEMOSTRACIÓN 1) Según el corolario del teorema A2, se tiene que

$$tr\ A^2 = \sum_{i=1}^{p} grad\ v_i$$

Basta entonces observar que cada una de las q líneas de G es incidente con 2 vértices, de modo que al sumar los grados de todos los vértices de G se estará considerando el doble de la cantidad de líneas de G.

2) Sea $A^3 = (c_{ij})_{i,j=1,\ldots,p}$. Según el teorema A2, c_{ii} = cantidad de caminos de longitud 3 que unen el vértice v_i consigo mismo. Obsérvese que cada ciclo de longitud 3 en G* en el que esté involucrado el vértice v_i corresponde con 2 caminos de longitud 3 que unen a v_i con v_i.

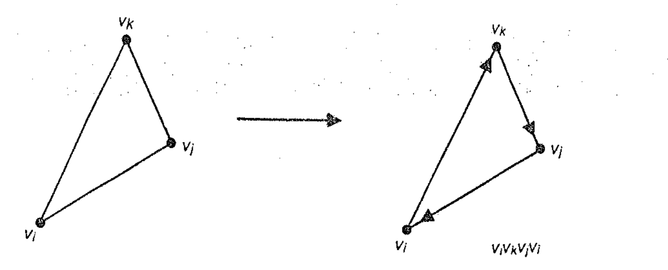


Figura 20

Como en cada ciclo de longitud 3 hay 3 vértices distintos de G involucrados, es claro entonces que al sumar la cantidad de caminos de longitud 3 que unen a cada uno de los vértices de G consigo mismo, se estará considerando 6 veces el número de ciclos de longitud 3 en G.

Q.E.D.

EJEMPLO 6

Por ejemplo, para la (5, 9)-gráfica de la figura 17 se tiene que

$$tr\ A^3 = 10 + 6 + 10 + 6 + 10 = 42$$

de modo que existen $\frac{1}{6}(42) = 7$ ciclos de longitud 3 en G. Éstos son:

$$v_1v_2v_3v_1,\ v_1v_3v_4v_1,\ v_4v_3v_5v_4,\ v_1v_4v_5v_1,\ v_2v_3v_5v_2,\ v_1v_3v_5v_1 \text{ y } v_1v_2v_5v_1.$$

*Dos ciclos en G se identifican si ambos involucran a los mismos vértices y tienen la misma longitud.

El resto de este apéndice se dedicará a estudiar el polinomio característico de una gráfica, así como algunas de las propiedades de sus raíces.

A.4. EL ESPECTRO DE UNA GRÁFICA

Ya se había observado anteriormente que es posible definir el *polinomio característico de una (p, q)-gráfica G*, como el polinomio característico de la matriz de adyacencia de alguna (de cualquiera, por el teorema A1 y el teorema A3 de este capítulo) de las posibles gráficas marcadas de G. Obsérvese que éste es un polinomio de grado p.

EJEMPLO 7 Por ejemplo, considere la (3, 2)-gráfica



Figura 21

Para obtener el polinomio característico de esta gráfica, marque de algún modo sus vértices como v_1, v_2, v_3, diga

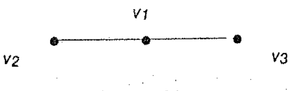


Figura 22

La matriz de adyacencia de esta gráfica es:

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Entonces el polinomio característico de G es:

$$p(\lambda) = det(A - \lambda I) = det \begin{bmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 0 \\ 1 & 0 & -\lambda \end{bmatrix} = -\lambda^3 + 2\lambda$$

Defínase el *espectro* de una (p, q)-gráfica G, denotado por E_G, como el conjunto de las p raíces de su polinomio característico. Es decir,

$$E_G = \{\lambda \mid p(\lambda) = 0\}$$

en donde $p(\lambda)$ es el polinomio característico de G.

TEOREMA A4 El espectro de una gráfica G es real (esto es, está constituido por números reales).

DEMOSTRACIÓN Se deduce inmediatamente del hecho de que una matriz de adyacencia de G es simétrica y del teorema 4.1 de este capítulo.

Q.E.D.

Por ejemplo, el espectro de la (3, 2)-gráfica de la figura 21 está constituido por las raíces de su polinomio característico $-\lambda^3 + 2\lambda$. Éste es entonces,

$$E_G = \{0, \sqrt{2}, -\sqrt{2}\}$$

A continuación se establecerán algunas propiedades muy interesantes que posee el espectro de una gráfica. Para hacerlo, se explotarán dos hechos ya estudiados en este capítulo. Éstos son:

a) Toda matriz simétrica es diagonalizable (teorema 4.3 de este capítulo).
b) Si A_1 y A_2 son matrices semejantes, su traza es la misma (ejercicio 5, sección 4, capítulo 4).

TEOREMA A5 Sea G una (p, q)-gráfica y sea

$$E_G = \{\lambda_1, \lambda_2, \ldots, \lambda_p\}$$

su espectro. Entonces,

1) $\lambda_1 + \lambda_2 + \ldots + \lambda_p = 0$
2) $\lambda_1^2 + \lambda_2^2 + \ldots + \lambda_p^2 = 2q$
3) $\lambda_1^3 + \lambda_2^3 + \ldots + \lambda_p^3 = 6C_3(G)$

En donde $C_3(G)$ es el número de ciclos de longitud 3 en G.

DEMOSTRACIÓN Sea A una matriz de adyacencia de la gráfica G. Como A es simétrica, A es semejante a la matriz diagonal

$$D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_p \end{bmatrix}$$

Más aún, A^n es semejante a la matriz

$$D = \begin{bmatrix} \lambda_1^n & & & 0 \\ & \lambda_2^n & & \\ & & \ddots & \\ 0 & & & \lambda_p^n \end{bmatrix}$$

para cada $n \in N$ (véase demostración del teorema 3.3 de este capítulo). Entonces,

$$tr\,A^n = tr\,D^n = \lambda_1^n + \lambda_2^n + \ldots + \lambda_p^n$$

para $n = 1$ se obtiene

$$\lambda_1 + \lambda_2 + \ldots + \lambda_p = tr\,A = 0$$

pues la matriz A tiene en su diagonal principal solamente ceros (ningún vértice de G es adyacente consigo mismo).

Para $n = 2$ se obtiene

$$\lambda_1^2 + \lambda_2^2 + \ldots + \lambda_p^2 = tr\,A^2 = 2q$$

por el inciso (1) del teorema A3.

Finalmente, para $n = 3$ se obtiene

$$\lambda_1^3 + \lambda_2^3 + \ldots + \lambda_p^3 = tr\,A^3 = 6C_3(G)$$

por el inciso (2) del mismo teorema A3.

Q.E.D.

Se desea resaltar el hecho de que las fórmulas establecidas en el teorema anterior, contienen en su primer miembro relaciones puramente algebraicas entre los elementos del espectro de la gráfica G, mientras que en el miembro derecho poseen información de la gráfica G misma. Es, pues, interesante observar que en ellas se establece una igualdad entre dos cantidades esencialmente distintas: una algebraica y otra geométrica.

EJEMPLO 8 Verifique el contenido del teorema anterior con la (3, 3)-gráfica

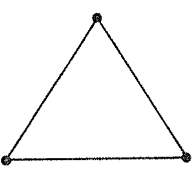


Figura 23

Marque sus vértices como v_1, v_2, v_3, dígase de la siguiente manera:

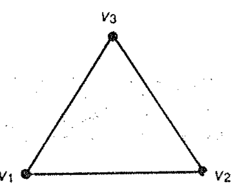


Figura 24

La matriz de adyacencia de esta gráfica es:

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

cuyo polinomio característico es:

$$p(\lambda) = det(A - \lambda I) = det\begin{bmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{bmatrix} = -\lambda^3 + 3\lambda + 2 = (\lambda + 1)^2(2 - \lambda)$$

Entonces el espectro de la gráfica G es:

$$E_G = \{-1, -1, 2\}$$

se comprueba pues que

$\lambda_1 + \lambda_2 + \lambda_3 = -1 - 1 + 2 = 0$

$\lambda_1^2 + \lambda_2^2 + \lambda_3^2 = (-1)^2 + (-1)^2 + (2)^2 = 6 = 2(3) = 2$ (número de lados de G).

$\lambda_1^3 + \lambda_2^3 + \lambda_3^3 = (-1)^3 + (-1)^3 + (2)^3 = 6 = 6(1) = 6$ (número de ciclos de longitud 3 de G).

A.5. UNA COTA PARA EL ESPECTRO

Se verá cómo la estructura geométrica de una gráfica determina cotas superior e inferior (en la recta real) para su espectro. El resultado que se quiere demostrar es el siguiente:

TEOREMA A6

Sea G una (p, q)-gráfica con vértices $v_1, v_2, \ldots, v_p$. Sea

$$D = \text{máx}\,\{grad\,v_i, \quad i = 1, 2, \ldots, p\}$$

(D es el máximo de los grados de los vértices de G). Si $E_G = \{\lambda_1, \lambda_2, \ldots, \lambda_p\}$ es el espectro de G, entonces

$$E_G \subset [-D, D]$$

Es decir que todos los elementos del espectro de G se encuentran en el intervalo cerrado $[-D, D]$.

Antes de ver la demostración de este resultado, véanse algunos ejemplos.

EJEMPLO 9 Para la (3, 2)-gráfica de la figura 22, se tiene que $grad\,v_2 = grad\,v_3 = 1$ y $grad\,v_1 = 2$ de modo que $D = 2$. Según el teorema anterior, el espectro de la gráfica debe estar contenido en el intervalo $[-2, 2]$. En efecto, se había visto que su espectro es $\{0, \sqrt{2}, -\sqrt{2}\}$.

Para la (3, 3)-gráfica de la figura 24 se tiene que $grad\,v_1 = grad\,v_2 = grad\,v_3 = 2$, de modo que $D = 2$. Nuevamente el teorema anterior asegura que el espectro de la gráfica debe encontrarse en el intervalo $[-2, 2]$. Se había ya encontrado que $E_G = \{-1, -1, 2\}$.

Como último ejemplo, se ve que la (6, 8)-gráfica G

Figura 25

tiene por polinomio característico a

$$p(\lambda) = \lambda^6 - 8\lambda^4 - 4\lambda^3 + 9\lambda^2 + 4\lambda - 1$$

En este caso se tiene que $D = grad\ v_5 = 4$. Según el teorema anterior, las 6 raíces (reales, por el teorema A4) de $p(\lambda)$ se encuentran en el intervalo $[-4, 4]$.

Es interesante notar de nuevo la combinación de informaciones algebraica y geométrica que establece este teorema.

La demostración del teorema A6 será prácticamente inmediata después de probar el siguiente resultado de carácter general:

TEOREMA A7

(TEOREMA DE LOS CÍRCULOS DE GERSHGORIN.) Sea A una matriz cuadrada de orden n con elementos a_{ij}, $i, j = 1, \ldots, n$. Sea $p(\lambda) = det(A - \lambda I)$ el polinomio característico de A. Si $\lambda_1, \lambda_2, \ldots, \lambda_n$ son las n raíces (complejas, en general) de $p(\lambda)$, entonces

$$\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \subset \bigcup_{i=1}^{n} D_i$$

en donde D_i es el círculo (en el plano complejo)

$$D_i = \{Z \in C \mid |Z - a_{ii}| \leq r_i\} \qquad r_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$$

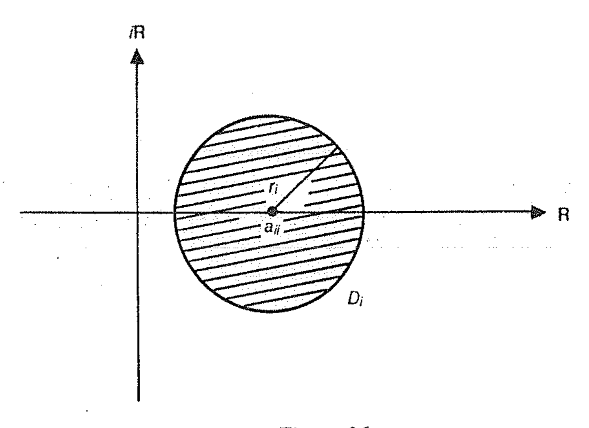


Figura 26

DEMOSTRACIÓN Sea λ un valor propio de A y $X = (x_1, x_2, \ldots, x_n)$ un vector propio asociado a λ. Entonces,

$$AX = \lambda X$$

Sea $i \in \{1, 2, \ldots, n\}$ tal que

$$|x_i| = \text{máx}\ \{|x_j|, j = 1, 2, \ldots, n\}$$

(es claro que $x_i \neq 0$, ¿por qué?). Al igualar los i-ésimos elementos de las matrices en la expresión $AX = \lambda X$ se obtiene

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = \lambda x_i$$

o bien,

$$(\lambda - a_{ii})x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j$$

Al tomar módulo de los números complejos en ambos miembros de esta última expresión se obtiene

$$|(\lambda - a_{ii})x_i| = \left| \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j \right|$$

Al usar las propiedades del módulo de números complejos* queda

$$|\lambda - a_{ii}|\,|x_i| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|\,|x_j|$$

o bien,

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \left|\frac{x_j}{x_i}\right| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| = r_i$$

en donde la última desigualdad queda justificada por la manera como fue escogido el índice i.

Esto muestra entonces que $\lambda \in D_i$ y por lo tanto,

$$\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \subset \bigcup_{i=1}^{n} D_i$$

Q.E.D.

EJEMPLO 10 Para ejemplificar el teorema anterior considere la matriz

$$A = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}$$

---

*Se usan las propiedades

1) $|Z_1 Z_2| = |Z_1|\,|Z_2|$

2) $|Z_1 + Z_2| \leq |Z_1| + |Z_2|$, $\quad Z_1, Z_2 \in C$

En este caso, el círculo D_1 del teorema es:

$$D_1 = \{Z \in \mathbf{C} \mid |Z - 1| \leq 1\}, \text{ pues } a_{11} = 1 \text{ y } r_1 = |-1| = 1$$

y el círculo D_2 es:

$$D_2 = \{Z \in \mathbf{C} \mid |Z + 1| \leq 2\}, \text{ pues } a_{22} = -1 \text{ y } r_2 = |2| = 2$$

Éstos se ven como

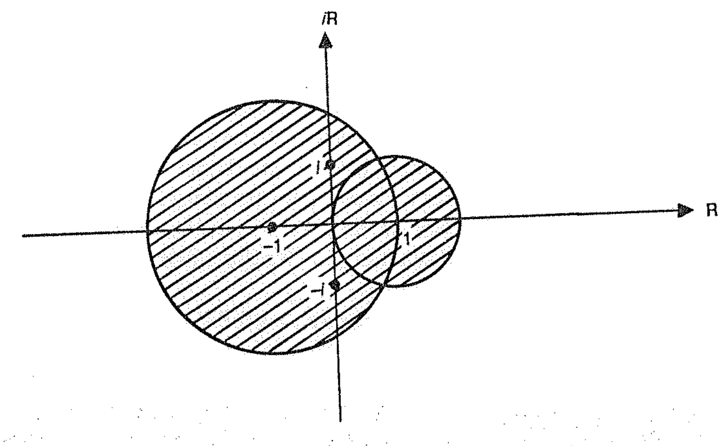


Figura 27

El teorema de los círculos de Gershgorin asegura que las raíces del polinomio característico de A se deben encontrar dentro de la región sombreada de la figura 27.

El polinomio característico de A es:

$$p(\lambda) = det(A - \lambda I) = det\begin{bmatrix} 1-\lambda & -1 \\ 2 & -1-\lambda \end{bmatrix} = \lambda^2 + 1$$

y sus raíces son $\lambda_{1,2} = \pm i$ las cuales efectivamente se encuentran en la unión de los círculos D_1 y D_2.

Al usar el teorema A7 se puede dar ahora un argumento muy simple que pruebe el teorema A6.

DEMOSTRACIÓN DEL TEOREMA A6

Si A es una matriz de adyacencia de la (p, q)-gráfica G, ésta tiene en su diagonal principal solamente ceros. Entonces, los círculos D_i del teorema A7 tendrán todos su centro en el origen. Obsérvese además que

$$r_i = \sum_{\substack{j=1 \\ j \neq i}} |a_{ij}| = grad\ v_i$$

por la misma construcción de la matriz A.

Se tiene entonces que los círculos D_i son todos concéntricos, siendo D_i de radio $grad\ v_i$. La unión de todos estos círculos es un nuevo círculo de centro en el origen y radio.

$$D = \text{máx}\ \{r_i, i = 1, 2, \ldots, p\} = \text{máx}\ \{grad\ v_i, i = 1, 2, \ldots, p\}$$

Entonces el espectro de G se debe encontrar limitado por este círculo. En particular, como el espectro es real, éste se encuentra en el intervalo $[-D, D]$, como se quería probar.

Q.E.D.

A.6. EL TEOREMA DE SACHS

En más de una ocasión en este apéndice, se ha visto cómo a partir de la geometría de una gráfica se ha obtenido información algebraica de la misma (de los elementos algebraicos asociados a la gráfica) y viceversa. El teorema de Sachs que se verá en este apartado, con el que se termina este apéndice, muestra de una manera magistral esta interrelación información geométrica-información algebraica que contiene una gráfica: en él se establecen fórmulas para los coeficientes del polinomio característico de una gráfica al usar solamente información geométrica de la misma.

Se introducirán primeramente algunos conceptos nuevos, que aparecen en el teorema de Sachs.

Sea G una (p, q)-gráfica. Una *subgráfica* de G es una gráfica que tiene todos sus vértices y lados en G.

EJEMPLO 11

Por ejemplo, considere la (4, 5)-gráfica G

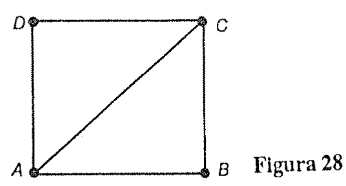
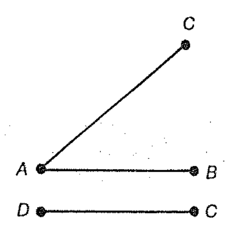


Figura 28

Las siguientes gráficas son ejemplos de subgráficas de G:



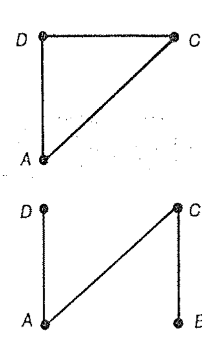


Figura 29

Sea G una (p, q)-gráfica y v_i un vértice de G. La *subgráfica* $G - v_i$ de G es aquella que tiene

a) todos los vértices de G excepto v_i
b) todos los lados que no son incidentes con v_i

Se dice que $G - v_i$ es una gráfica que se ha obtenido de G eliminando en ésta el vértice v_i.

EJEMPLO 12 Por ejemplo, las siguientes 4 subgráficas de la gráfica de la figura 28, se han obtenido eliminando en ésta los vértices A, B, C y D, respectivamente.

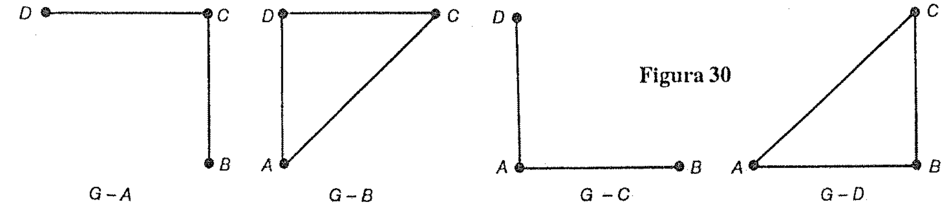


Figura 30

Similarmente, las siguientes subgráficas de G (Figura 28) se han obtenido eliminando dos vértices de ésta:

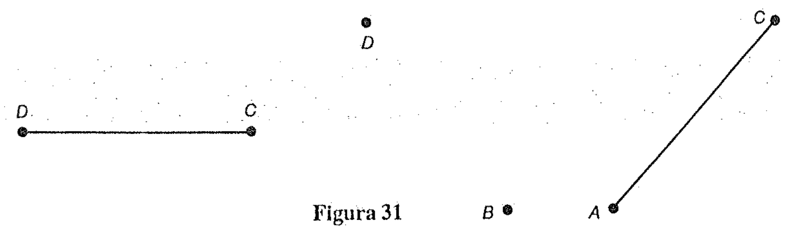


Figura 31

Dada una (p, q)-gráfica G, se estará interesado en considerar subgráficas de G con k vértices, $k \leq p$, y que estén constituidas solamente por

1) ciclos

y/o

2) fragmentos del tipo •———•

Se llamarán a estas subgráficas de G, *subgráficas de Sachs de G con k vértices.*

EJEMPLO 13 Por ejemplo, las subgráficas de Sachs de la (4, 5)-gráfica G de la figura 28, con 4 vértices son

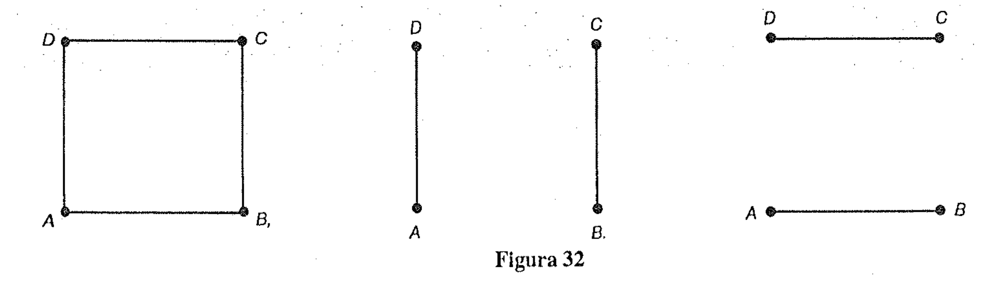


Figura 32

mientras que las subgráficas de Sachs de G con 3 vértices son:

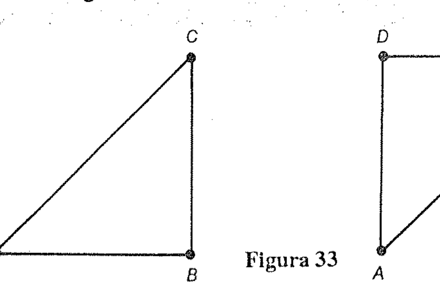


Figura 33

Sea G_i una subgráfica de Sachs de una gráfica G. Se denotará por c_i al número de componentes de G_i (número de piezas que constituyen a G_i) y por e_i al número de ciclos de G_i (no importa de qué longitud).

Por ejemplo, para las subgráficas de Sachs de la figura 32 se tiene, para la primera de ellas, $c_1 = 1$, $e_1 = 1$, para las dos últimas $c_2 = c_3 = 2$, $e_2 = e_3 = 0$. Similarmente, para las dos subgráficas de Sachs de la figura 33, se tiene $c_1 = c_2 = 1$, $e_1 = e_2 = 1$.

Se está ahora en posibilidades de enunciar el teorema de Sachs. Su demostración está fuera de los objetivos de este apéndice.

TEOREMA A7 (TEOREMA DE SACHS.) Sea G una (p, q)-gráfica y sea

$$p(\lambda) = (-1)^n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$$

su polinomio característico. Los coeficientes $a_0, a_1, \ldots, a_{n-1}$ de $p(\lambda)$ pueden ser calculados por las fórmulas

$$a_{n-r} = (-1)^n \sum_{S_r} (-1)^{c_i}(2)^{e_i} \qquad 1 \leq r \leq n$$

en donde S_r es el conjunto de subgráficas de Sachs de G con r vértices y la suma se efectúa sobre todos los elementos de este conjunto, cuyo i-ésimo elemento tiene c_i componentes y e_i ciclos. Si S_r es vacío, $a_{n-r} = 0$.

Obsérvese que siempre se tendrá $a_{n-1} = 0$, pues el conjunto de subgráficas de Sachs de una gráfica G, con 1 vértice es vacío.

EJEMPLO 14 Como un primer ejemplo, obtenga el polinomio característico de la (4, 5)-gráfica G de la figura 28.

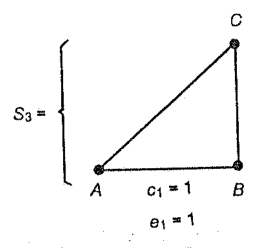
Para calcular a_0 ($r = 4$), se obtienen las subgráficas de Sachs de G con 4 vértices. Como ya se había visto, éstas son las gráficas de la figura 32. Entonces,

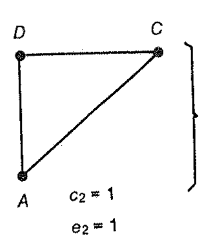
$S_4 =$ D C A B $c_1 = 1$ $e_1 = 1$; D C A B $c_2 = 2$ $e_2 = 0$; D C A B $c_3 = 2$ $e_3 = 0$

Entonces,

$$a_0 = (-1)^4 [(-1)^1(2)^1 + (-1)^2(2)^0(2)] = 0$$

Se calcula ahora $a_1(r = 3)$. En este caso, S_3 está formado por las gráficas de la figura 33.

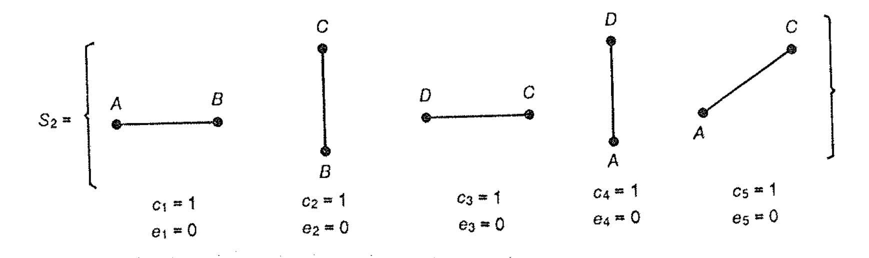




Entonces,

$$a_1 = (-1)^4 [(-1)^1(2)^2(2)] = -4$$

Calcúlese por último $a_2(r = 2)$. Obsérvese que cada lado de G es una subgráfica de Sachs con dos vértices. Entonces,



Entonces,

$$a_2 = (-1)^4 [(-1)^2(2)^0(5)] = -5$$

y por lo tanto, el polinomio característico de G es:

$$p(\lambda) = \lambda^4 - 5\lambda^2 - 4\lambda$$

como puede verificarse directamente.

EJEMPLO 15 Vea otro ejemplo. Considere la (5, 7)-gráfica G.

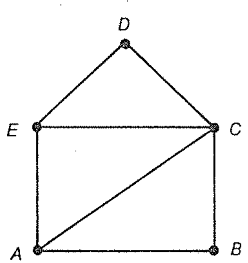
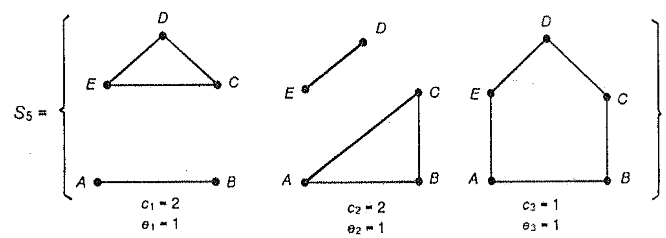


Figura 34

y obtenga su polinomio característico usando el teorema de Sachs.

Cálculo de a_0: En este caso S_5 es el conjunto

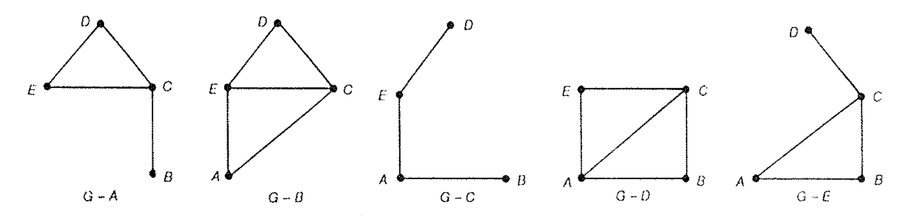


y por lo tanto,

$$a_0 = (-1)^5 [(-1)^2(2)^1(2) + (-1)^1(2)^1] = -2$$

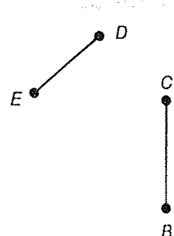
Cálculo de a_1: Se tienen que considerar ahora todas las subgráficas de Sachs de G con 4 vértices. Se puede proceder sistemáticamente obteniendo primero las subgráficas de G con 4 vértices (obtenidas eliminando de G uno de sus vértices) y obteniendo luego las subgráficas de Sachs de cada una de estas subgráficas.

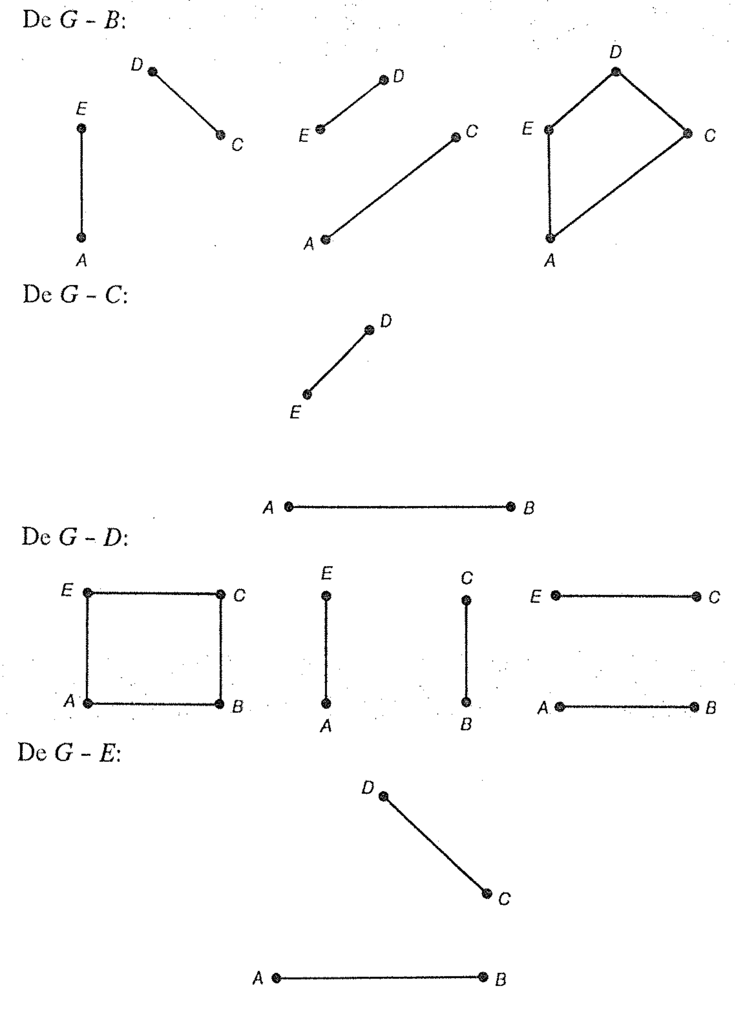
Las subgráficas de G con 4 vértices son:



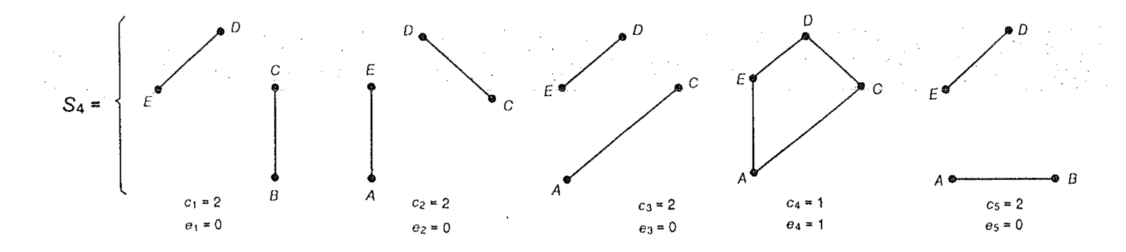
De cada una de ellas considere sus subgráficas de Sachs

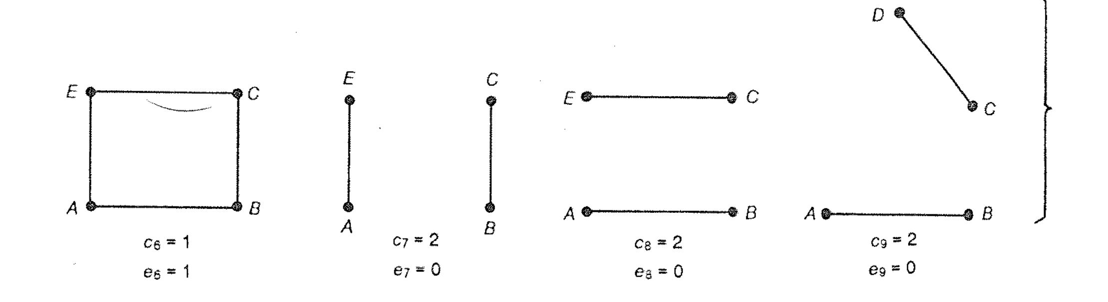
De $G - A$:





Entonces,

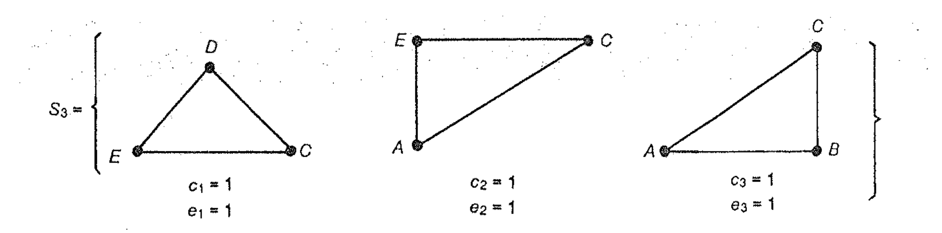




y por lo tanto,

$$a_1 = (-1)^5 [(-1)^2(2)^0(7) + (-1)^1(2)^1(2)] = -3$$

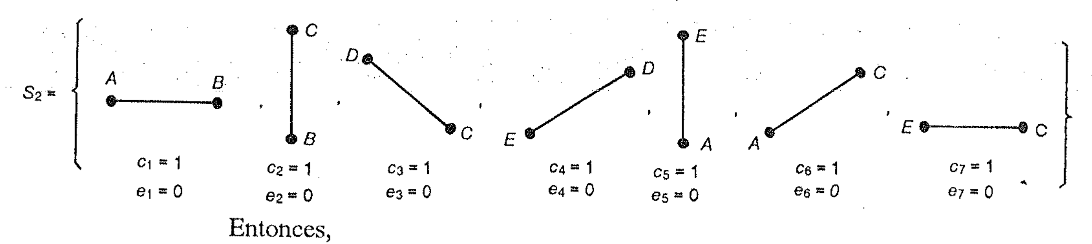
Cálculo de a_2: No es difícil ver que en este caso el conjunto de subgráficas de Sachs de G con 3 vértices es:



Entonces,

$$a_2 = (-1)^5 [(-1)^1(2)^1(3)] = 6$$

Cálculo de a_3: Nuevamente se observa que las subgráficas de Sachs de G con 2 vértices son precisamente los lados de G; entonces



Entonces,

$$a_3 = (-1)^5 [(-1)^1(2)^0(7)] = 7$$

y por lo tanto, el polinomio característico de G es:

$$p(\lambda) = -\lambda^5 + 7\lambda^3 + 6\lambda^2 - 3\lambda - 2$$

EJERCICIOS (APÉNDICE, CAPÍTULO 6)

1. Dibuje una representación geométrica para cada una de las siguientes gráficas, de las que se dan el conjunto V de vértices y el conjunto X de lados

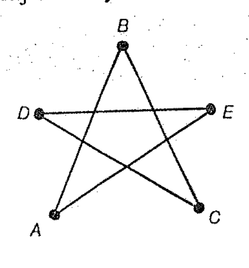
   a) $V = \{A, B, C\}$
   $X = \{(A, B), (B, C), (C, A)\}$

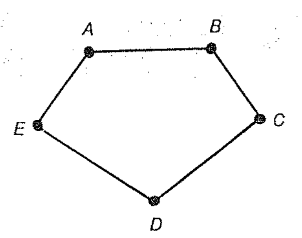
   b) $V = \{A, B, C, D\}$
   $X = \{(A, B), (B, C), (A, D)\}$

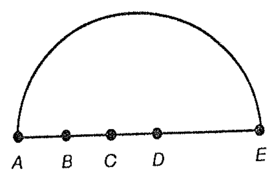
   c) $V = \{A, B, C, D, E, F\}$
   $X = \{(A, B), (B, C), (C, A), (D, E), (D, F)\}$

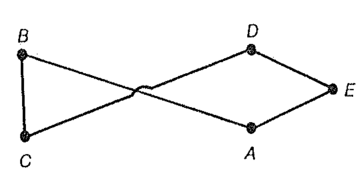
   d) $V = \{A, B, C, D, E, F, G, H, I\}$
   $X = \{(A, B), (A, C), (B, C), (C, D), (D, E), (D, F), (F, G), (F, H), (H, I)\}$

2. ¿Cuáles de los siguientes diagramas representan a una misma gráfica G? Describa los conjuntos V y X de G

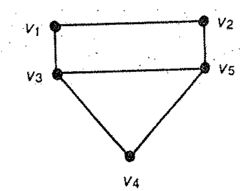


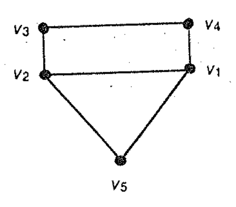






3. Considere las siguientes dos posibilidades de marcar la gráfica mostrada:

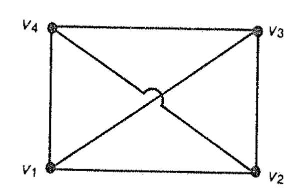


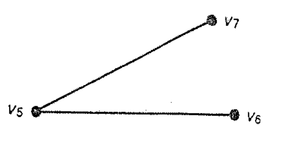


   a) Obtenga la matriz de adyacencia de cada una de estas gráficas.
   b) Verifique la validez del teorema A1 en este caso particular.

4. Sea G una (p, q)-gráfica no conexa constituida por dos componentes, dígase G_1 con p_1 vértices y G_2 con p_2 vértices $(p_1 + p_2 = q)$. Describa la estructura de la matriz de adyacencia de G.

5. Considere la siguiente (7, 8)-gráfica no conexa:

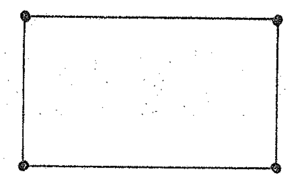




   Calcule el número de:

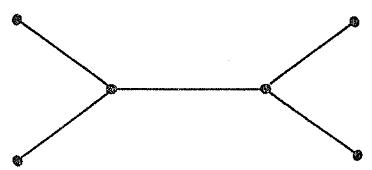
   a) $(v_1 - v_3)$ - caminos de longitud 2.
   b) $(v_1 - v_7)$ - caminos de longitud 2.
   c) $(v_2 - v_3)$ - caminos de longitud 2.
   d) $(v_5 - v_7)$ - caminos de longitud 3.
   e) $(v_4 - v_6)$ - caminos de longitud 3.
   f) $(v_3 - v_4)$ - caminos de longitud 4.
   g) $(v_3 - v_5)$ - caminos de longitud 4.
   h) $(v_7 - v_6)$ - caminos de longitud 4.
   i) $(v_1 - v_2)$ - caminos de longitud 5.
   j) $(v_1 - v_4)$ - caminos de longitud 5.
   k) $(v_5 - v_6)$ - caminos de longitud 5.
   l) $(v_4 - v_6)$ - caminos de longitud 5.

6. Para la siguiente (4, 4)-gráfica G:

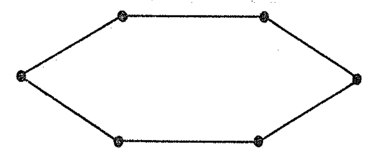


   a) Determine su polinomio característico.
   b) Encuentre el intervalo $[-D, D]$ en el que se debe encontrar su espectro.
   c) Determine su espectro.
   d) Verifique la validez del teorema A4.

7. Repita el ejercicio anterior con la (6, 5)-gráfica

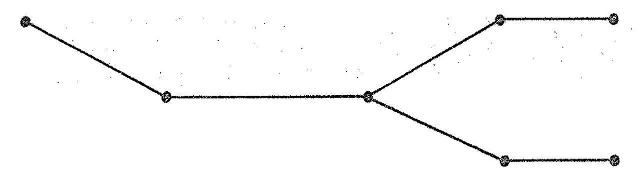


8. Use el teorema de Sachs para hallar el polinomio característico de la (6, 6)-gráfica



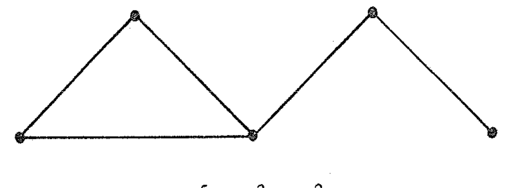
   Determine el espectro de esta gráfica y verifique la validez de los teoremas A5 y A6.

9. Use el teorema de Sachs para hallar el polinomio característico de la (7, 6)-gráfica



Determine también su espectro y verifique la validez de los teoremas A5 y A6.

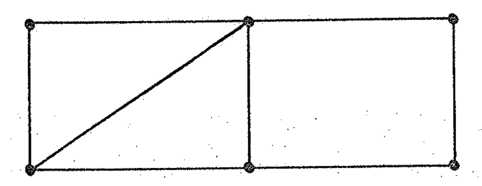
10. Use el teorema de Sachs para demostrar que el polinomio característico de la (5, 5)-gráfica



es:

$$p(\lambda) = -\lambda^5 + 5\lambda^3 + 2\lambda^2 - 4\lambda - 2$$

11. Use el teorema de Sachs para demostrar que el polinomio característico de la (6, 8)-gráfica



es:

$$p(\lambda) = \lambda^6 - 8\lambda^4 - 4\lambda^3 + 9\lambda^2 + 4\lambda - 1$$

② 12. Dos gráficas G_1 y G_2 se dicen ser *isoespectrales* si ellas poseen el mismo polinomio característico (y por lo tanto, el mismo espectro). Demuestre que las siguientes (10, 10)-gráficas son isoespectrales



① 13. En este apéndice aparecieron los dos siguientes hechos:

a) La traza de cualquier matriz de adyacencia de una (p, q)-gráfica G es cero.

b) El conjunto de subgráficas de Sachs de una (p, q)-gráfica G, con un solo vértice, es vacío.

Discuta la relación entre estos hechos.

(*Sugerencia*: ambos tienen que ver con el coeficiente de λ^{p-1} en el polinomio característico de G.)

CAPÍTULO SIETE

Formas bilineales y cuadráticas

En el capítulo 4 se estudió el concepto de linealidad para una función $f: U \to V$ entre los espacios vectoriales U y V. Una generalización de la teoría ahí desarrollada va en la dirección de estudiar funciones de varias variables que conservan la propiedad de linealidad en cada una de ellas; es decir, funciones del tipo

$$f: U_1 \times U_2 \times \ldots \times U_n \to V$$

($U_1, U_2, \ldots, U_n$ y V espacios vectoriales) tal que para cada $i = 1, 2, \ldots, n$ se tiene la siguiente propiedad: si se fijan $n - 1$ vectores $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n$ en los espacios $U_1, \ldots, U_{i-1}, U_{i+1}, \ldots, U_n$, respectivamente, la función

$$\phi: U_i \to V$$

$$\phi(u_i) = f(u_1, \ldots, u_{i-1}, u_i, u_{i+1}, \ldots, u_n)$$

es una transformación lineal entre los espacios U_i y V.

A las funciones de este tipo se les llama *funciones multilineales* (cuando $V = \mathbf{R}$ se les llama *formas multilineales*) y a la parte del álgebra que se encarga de su estudio se le llama *álgebra multilineal.*

En este capítulo será de interés estudiar las *formas bilineales* (en la notación anterior, $n = 2$, y $V = \mathbf{R}$) y otro tipo de funciones procedentes de ellas llamadas *formas cuadráticas*. Aunque éstas resulten casos muy particulares del tipo de funciones de las que se encarga el álgebra multilineal, resultan ser de gran utilidad práctica en muchas otras partes de la matemática. En este mismo capítulo (secciones 3 y 4) se verá que con la herramienta desarrollada en las secciones 1 y 2 se podrá resolver un problema de clasificación de lugares geométricos de curvas en el plano y superficies en el espacio.

1. FORMAS BILINEALES

En el capítulo 4 se introdujo el concepto de linealidad de una función entre dos espacios vectoriales (que se llama transformación lineal). Más concretamente, en el apéndice 1 de tal capítulo se estudió un tipo de transformaciones lineales cuyo

codominio era el conjunto de números reales, que se llaman funcionales lineales. Este tipo de funciones tenía un solo argumento (una sola variable) que era elemento de un cierto espacio vectorial. Si se consideran ahora funciones con dos argumentos (con dos variables) también tomando valores en los reales, tales que fijando uno de sus argumentos se obtiene una función lineal (un funcional lineal) respecto del otro, se obtiene una "forma bilineal".

Más precisamente, se tiene la siguiente definición:

DEFINICIÓN 1.1 Sea V un espacio vectorial. Una *forma bilineal* (real) en V es una función $f: V \times V \to \mathbf{R}$ que asocia a cada par de vectores $v_1, v_2 \in V$ un número real $f(v_1, v_2)$ que satisface

$$f(cv_1 + v_1', v_2) = cf(v_1, v_2) + f(v_1', v_2) \qquad (1.1)$$

$$f(v_1, cv_2 + v_2') = cf(v_1, v_2) + f(v_1, v_2'), \quad c \in \mathbf{R}$$

Cada una de las expresiones en (1.1) en la definición anterior, establece la condición de *linealidad* de la forma bilineal f respecto de cada una de sus dos variables cuando la otra se mantiene fija.

En efecto, si se fija el vector $v_2 \in V$ (la segunda variable de f) y se considera la función

$$\phi: V \to \mathbf{R}$$

$$\phi(v) = f(v, v_2)$$

se obtiene que

$$\phi(cv_1 + v_1') = f(cv_1 + v_1', v_2) = cf(v_1, v_2) + f(v_1', v_2)$$

$$= c\phi(v_1) + \phi(v_1')$$

lo que muestra que ϕ es lineal y así entonces se dice que "f es lineal respecto de su primera variable".

Similarmente se puede ver que la función

$$\psi: V \to \mathbf{R}$$

$$\psi(v) = f(v_1, v) \qquad v_1 \in V, \text{ fijo}$$

es una función lineal, lo que muestra que f es lineal respecto de su segunda variable.

Véanse algunos ejemplos.

EJEMPLO 1 Si V es un espacio vectorial en el que está definido el producto interno $(\cdot \mid \cdot)$, este producto interno es una forma bilineal en V. Más precisamente, la función $f: V \times V \to \mathbf{R}$ dada por

$$f(v_1, v_2) = (v_1 \mid v_2)$$

es una forma bilineal en V. Esto se desprende fácilmente de la definición de producto interno dada en la sección 1 del capítulo 5.

El siguiente ejemplo es un ejemplo muy importante de formas bilineales, pues más adelante se verá que toda forma bilineal en un espacio vectorial de dimensión finita puede verse como una forma del tipo que ahora se describirá.

EJEMPLO 2 Sea $V = M_{n \times 1}$ y sea A una matriz cuadrada de orden n fija. La función $f: M_{n \times 1} \times M_{n \times 1} \to \mathbf{R}$ dada por

$$f(X, Y) = X^tAY \qquad (1.2)$$

(en donde se está identificando la matriz de orden 1 $X^tAY = [a]$ con el número real a) es una forma bilineal en $M_{n \times 1}$.

En efecto, si c es cualquier escalar se tiene:

$$f(cX_1 + X_2, Y) = (cX_1 + X_2)^tAY = (cX_1^t + X_2^t)AY = cX_1^tAY + X_2^tAY$$

$$= cf(X_1, Y) + f(X_2, Y)$$

$$f(X, cY_1 + Y_2) = X^tA(cY_1 + Y_2) = X^tAcY_1 + X^tAY_2 = cX^tAY_1 + cX^tAY_2$$

$$= cf(X, Y_1) + f(X, Y_2)$$

Por ejemplo, si $n = 2$ y A es la matriz

$$A = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}$$

la forma bilineal del ejemplo anterior sería:

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = [x_1\ x_2] \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1y_1 + 3x_1y_2 - x_2y_1 + 2x_2y_2$$

Suponga ahora que se tiene la forma bilineal en $M_{2 \times 1}$

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = 2x_1y_1 + 4x_1y_2 - 8x_2y_1 + 3x_2y_2$$

No es difícil darse cuenta que esta forma puede ser escrita como

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = [x_1\ x_2] \begin{bmatrix} 2 & 4 \\ -8 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

que es una forma del tipo descrito en (1.2).

Más generalmente, sea V un espacio vectorial de dimensión 2 y sea $\beta = \{v_1, v_2\}$ una base de V. Vea cómo describir la estructura general que tiene una forma bilineal f en V.

Dados $x, y \in V$, existen escalares x_1, x_2, y_1, y_2 tales que

$$x = x_1v_1 + x_2v_2$$
$$y = y_1v_1 + y_2v_2$$

Al aplicar las condiciones de linealidad para f establecidas en (1.1) se obtiene

$$\begin{aligned} f(x, y) &= f(x_1v_1 + x_2v_2, y_1v_1 + y_2v_2) \\ &= f(v_1, v_1)x_1y_1 + f(v_1, v_2)x_1y_2 + f(v_2, v_1)x_2y_1 + f(v_2, v_2)x_2y_2 \end{aligned}$$

Se escribe

$$a_{ij} = f(v_i, v_j) \qquad i, j = 1, 2$$

Entonces la forma f queda escrita como

$$f(x, y) = a_{11}x_1y_1 + a_{12}x_1y_2 + a_{21}x_2y_1 + a_{22}x_2y_2$$

o bien,

$$\begin{aligned} f(x, y) &= [x_1\ x_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= [x]_\beta^t\ A[y]_\beta \end{aligned}$$

en donde $A = (f(v_i, v_j))_{i,j=1,2}$.

Se ha podido establecer entonces que la estructura general de una forma bilineal f en un espacio vectorial de dimensión 2 puede verse, por medio de una base β de este espacio, como una forma bilineal del tipo descrito en (1.2).

El objetivo de la siguiente subsección es llevar esta discusión al caso general de un espacio vectorial de dimensión n.

1.1. LA MATRIZ ASOCIADA A UNA FORMA BILINEAL

A partir de este momento y por lo que resta de la presente sección, se centrará la atención a estudiar formas bilineales en espacios vectoriales de *dimensión finita*.

Sea V un espacio vectorial de dimensión n. Fíjese una base $\beta = \{v_1, v_2, \ldots, v_n\}$ de V. Sea $f\colon V \times V \to \mathbf{R}$ una forma bilineal en V. Para dos vectores $x, y \in V$ se tiene

$$x = x_1v_1 + x_2v_2 + \ldots + x_nv_n = \sum_{i=1}^{n} x_iv_i$$

$$y = y_1v_1 + y_2v_2 + \ldots + y_nv_n = \sum_{j=1}^{n} y_jv_j$$

Entonces,

$$f(x, y) = f\left(\sum_{i=1}^{n} x_iv_i, \sum_{j=1}^{n} y_jv_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} f(v_i, v_j)x_iy_j$$

en donde la última igualdad se obtiene usando repetidas veces las condiciones (1.1) establecidas en la definición 1.1.

Escríbase $a_{ij} = f(v_i, v_j)$, $i, j = 1, 2, \ldots, n$. Sea $A = (a_{ij})_{i,j=1,\ldots,n}$. Se tiene entonces que

$$f(x, y) = \sum_{i,j=1}^{n} a_{ij}x_iy_j$$

Pero

$$[x]_\beta^t A[y]_\beta = \sum_{i,j=1}^{n} a_{ij}x_iy_j = f(x, y)$$

de modo que

$$f(x, y) = [x]_\beta^t A[y]_\beta \tag{1.3}$$

A la matriz A se le llama *matriz de la forma bilineal f respecto de la base* β, la cual se denota por $[f]_\beta$.

Se ha asociado entonces a la forma bilineal f en el espacio vectorial V de dimensión n, una matriz cuadrada A de orden n de modo que vale la fórmula (1.3). En este punto es conveniente hacer un par de observaciones importantes.

Primeramente obsérvese que los elementos de la matriz A *dependen* de la base β que se fije en V: no se puede hablar entonces de *la* matriz asociada a una forma bilineal f en V (una situación similar acontece con las matrices asociadas a transformaciones lineales). Se debe referir a la matriz de la forma f *respecto de la base* β.

La segunda observación que se debe hacer es que la matriz A en la expresión (1.3) depende única y exclusivamente de la base que se fije en el espacio vectorial V y *no* depende de los vectores X y Y que aparecen en tal expresión.

Como un corolario de las dos observaciones anteriores se puede concluir que, una vez fijada la base β del espacio vectorial V, la matriz asociada a la *forma bilineal* $f\colon V \times V \to \mathbf{R}$ respecto de β es *única*.

EJEMPLO 3 Por ejemplo, si V es el espacio $\mathbf{R}^3$ y $x = (x_1, x_2, x_3)$ y $y = (y_1, y_2, y_3)$, considere la forma bilineal $f\colon \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$ dada por

$$f(x, y) = x_1y_1 + 2x_2y_2 + x_2y_1 - 3x_2y_3 + x_3y_2 + 3x_3y_3 \tag{1.4}$$

Si $\beta_1 = \{e_1, e_2, e_3\}$ es la base canónica de $\mathbf{R}^3$, se tiene que

$$f(e_1, e_1) = 1 \quad f(e_2, e_1) = 1 \quad f(e_3, e_1) = 0$$
$$f(e_1, e_2) = 0 \quad f(e_2, e_2) = 2 \quad f(e_3, e_2) = 1$$
$$f(e_1, e_3) = 0 \quad f(e_2, e_3) = -3 \quad f(e_3, e_3) = 3$$

y entonces, la matriz asociada a f respecto de la base canónica de $\mathbf{R}^3$ es:

$$[f]_{\beta_1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & -3 \\ 0 & 1 & 3 \end{bmatrix}$$

de modo que

$$f(x, y) = [x_1\, x_2\, x_3] \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & -3 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Por ejemplo, si $x = (3, -2, 8)$, $y = (3, 4, -1)$, se tiene que

$$f(x, y) = [3\ -2\ 8] \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & -3 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ -1 \end{bmatrix} = -11$$

como puede comprobarse directamente con la fórmula (1.4).

Si se toma ahora la base $\beta_2 = \{v_1, v_2, v_3\}$ de $\mathbf{R}^3$ en donde

$$v_1 = (1, 0, 1)$$
$$v_2 = (0, -1, 2)$$
$$v_3 = (1, 1, 0)$$

Se tiene que

$$f(v_1, v_1) = 4 \quad f(v_2, v_1) = 8 \quad f(v_3, v_1) = -1$$
$$f(v_1, v_2) = 5 \quad f(v_2, v_2) = 18 \quad f(v_3, v_2) = -8$$
$$f(v_1, v_3) = 2 \quad f(v_2, v_3) = -1 \quad f(v_3, v_3) = 4$$

de modo que

$$[f]_{\beta_2} = \begin{bmatrix} 4 & 5 & 2 \\ 8 & 18 & -1 \\ -1 & -8 & 4 \end{bmatrix}$$

Si se toman nuevamente los vectores $x = (3, -2, 8)$ y $y = (3, 4, -1)$ en $\mathbf{R}^3$ se observa que

$$[x]_{\beta_2} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \qquad [y]_{\beta_2} = \begin{bmatrix} -1 \\ 0 \\ 4 \end{bmatrix}$$

se observa que, según la fórmula (1.3) se tiene

$$f(x, y) = [x]^t_{\beta_2} [f]_{\beta_2} [y]_{\beta_2} = [2\ 3\ 1] \begin{bmatrix} 4 & 5 & 2 \\ 8 & 18 & -1 \\ -1 & -8 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 4 \end{bmatrix} = -11$$

valor que ya se había calculado con la matriz asociada a f respecto de la base canónica de $\mathbf{R}^3$.

Resulta natural esperar que las matrices $[f]_{\beta_1}$ y $[f]_{\beta_2}$ del ejemplo anterior estén relacionadas entre sí de alguna manera, pues finalmente ellas representan a la misma forma bilineal f —en diferentes bases de $\mathbf{R}^3$—. El objetivo de la siguiente subsección es analizar esta relación.

1.2. CAMBIO DE BASE. RANGO DE UNA FORMA BILINEAL

Sea $f\colon V \times V \to \mathbf{R}$ una forma bilineal en el espacio vectorial V de dimensión n.

Sea β_1 una base V. En la subsección anterior se estableció la fórmula

$$f(x, y) = [x]^t_{\beta_1} A [y]_{\beta_1}$$

en donde $A = [f]_{\beta_1} = (f(v_i, v_j))_{i,j=1,\ldots,n}$ es la matriz de la forma f respecto de la base $\beta_1 = \{v_1, v_2, \ldots, v_n\}$.

Considérese otra base $\beta_2 = \{u_1, u_2, \ldots, u_n\}$ del espacio V. En la sección 6 (subsección 6.1) del capítulo 3 se relacionó la matriz de coordenadas de un vector $w \in V$ respecto de la base β_1 con la matriz de coordenadas de w respecto de la base β_2. El resultado fue

$$[w]_{\beta_1} = P[w]_{\beta_2}$$

en donde P es una matriz inversible de orden n: la matriz de cambio de base de β_2 a β_1 (la j-ésima columna de P está constituida por las coordenadas del j-ésimo vector de la base β_2 respecto de la base β_1).

Entonces, para $X, Y \in V$ se tiene

$$[x]_{\beta_1} = P[x]_{\beta_2}$$
$$[y]_{\beta_1} = P[y]_{\beta_2}$$

de modo que sustituyendo en (1.3) se obtiene

$$f(x, y) = (P[x]_{\beta_2})^t A(P[y]_{\beta_2}) = [x]^t_{\beta_2}(P^tAP)[y]_{\beta_2}$$

Debido a la unicidad de la matriz asociada a una forma bilineal respecto de una base dada del espacio V, se concluye de la expresión anterior que la matriz asociada a la forma bilineal f respecto de la base β_2 de V es P^tAP.

En resumen, se ha probado el siguiente resultado:

TEOREMA 1.1

Sea $f: V \times V \to \mathbf{R}$ una forma bilineal en el espacio vectorial V de dimensión n. Sean β_1 y β_2 dos bases de V. Si $[f]_{\beta_i}$ es la matriz asociada a la forma f respecto de la base β_i, $i = 1, 2$, se tiene

$$[f]_{\beta_2} = P^t[f]_{\beta_1}P \qquad (1.5)$$

En donde P es la matriz (inversible de orden n) de cambio de base de β_2 a β_1.

EJEMPLO 4

Por ejemplo, si se considera la forma bilineal f en $\mathbf{R}^3$ dada por (1.5), se tiene, como ya se había visto, que la matriz asociada a esta forma respecto de la base canónica $\beta_1 = (e_1, e_2, e_3)$ de $\mathbf{R}^3$ es:

$$A = [f]_{\beta_1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & -3 \\ 0 & 1 & 3 \end{bmatrix}$$

Si se toma ahora la base $\beta_2 = \{(1, 0, 1), (0, -1, 2), (1, 1, 0)\}$, siendo P la matriz de cambio de base de β_2 a β_1, se tiene que según el teorema 1.1, la matriz asociada a la forma f respecto de la base β_2 es:

$$[f]_{\beta_2} = P^tAP = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 2 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & -3 \\ 0 & 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 1 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 2 \\ 8 & 18 & -1 \\ -1 & -8 & 4 \end{bmatrix}$$

resultado que ya se había establecido anteriormente calculando directamente los elementos de esta matriz.

Tome ahora una forma bilineal $f: V \times V \to \mathbf{R}$ en el espacio vectorial V y dos bases distintas β_1 y β_2 de V. Según el ejercicio 10 de la sección 7 capítulo 4, se tiene que el rango de la matriz $[f]_{\beta_1}$ es el mismo que el rango de la matriz $[f]_{\beta_2}$ pues la matriz P de cambio de base de β_2 a β_1 que aparece en la fórmula (1.5) es una matriz inversible. Este hecho permite establecer la siguiente definición:

DEFINICIÓN 1.2

Sea $f: V \times V \to \mathbf{R}$ una forma bilineal en el espacio vectorial V de dimensión n. Se llama *rango* de la forma f al rango de la matriz asociada a f respecto de alguna (cualquiera) base de V.

EJEMPLO 5

Así por ejemplo, la forma bilineal en $\mathbf{R}^3$ de la fórmula (1.4) tiene rango 3, pues éste es el rango de la matriz asociada a f respecto de la base canónica β_1 de $\mathbf{R}^3$

$$[f]_{\beta_1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & -3 \\ 0 & 1 & 3 \end{bmatrix}$$

que es el mismo rango de la matriz asociada a f respecto de la base $\beta_2 = \{(1, 0, -1), (0, -1, 2), (1, 1, 0)\}$

$$[f]_{\beta_2} = \begin{bmatrix} 4 & 5 & 2 \\ 8 & 18 & -1 \\ -1 & -8 & 4 \end{bmatrix}$$

Una forma bilineal de $f: V \times V \to \mathbf{R}$ se dice ser *no degenerada* si su rango es igual a la dimensión del espacio V. Caso contrario, se dice que la forma f es *degenerada*.

La forma dada por la fórmula (1.4) es un ejemplo de forma bilineal no degenerada.

1.3. EL ESPACIO DE FORMAS BILINEALES

Sea V un espacio vectorial de dimensión n. Considérese el conjunto de todas las formas bilineales en V, al cual se denotará por $L_2(V, \mathbf{R})$.

Se define la suma y producto por escalares en $L_2(V, \mathbf{R})$ de la manera natural.

$$f_1, f_2 \in L_2(V, \mathbf{R}), f_1 + f_2: V \times V \to \mathbf{R}$$
$$(f_1 + f_2)(x, y) = f_1(x, y) + f_2(x, y)$$
$$f \in L_2(V, \mathbf{R}),\ c \in \mathbf{R},\ cf: V \times V \to \mathbf{R}$$
$$(cf)(x, y) = cf(x, y)$$

Se afirma que con la definición anterior se tiene $f_1 + f_2 \in L_2(V, \mathbf{R})$ y $cf \in L_2(V, \mathbf{R})$.

En efecto, si $x, x', y, y' \in V$ y $c \in \mathbf{R}$ se tiene

$$\begin{aligned}(f_1 + f_2)(cx + x', y) &= f_1(cx + x', y) + f_2(cx + x', y) \\ &= cf_1(x, y) + f_1(x', y) + cf_2(x, y) + f_2(x', y) \\ &= c(f_1(x, y) + f_2(x, y)) + (f_1(x', y) + f_2(x', y)) \\ &= c(f_1 + f_2)(x, y) + (f_1 + f_2)(x', y)\end{aligned}$$

lo que muestra que $f_1 + f_2$ es lineal respecto de su primera variable. Similarmente

$$\begin{aligned}(f_1 + f_2)(x, y + cy') &= f_1(x, y + cy') + f_2(x, y + cy')\\ &= f_1(x, y) + cf_1(x, y') + f_2(x, y) + cf_2(x, y')\\ &= f_1(x, y) + f_2(x, y) + c(f_1(x, y') + f_2(x, y'))\\ &= (f_1 + f_2)(x, y) + c(f_1 + f_2)(x, y')\end{aligned}$$

lo que muestra que $f_1 + f_2$ es lineal respecto de su segunda variable. Entonces, $f_1 + f_2 \in L_2(V, \mathbf{R})$.

La verificación de que $cf \in L_2(V, \mathbf{R})$ se deja como ejercicio para el lector.

Es también un simple ejercicio que se deja para el lector, verificar que el conjunto $L_2(V, \mathbf{R})$ con las operaciones de suma y producto por escalares anteriormente definidas es un espacio vectorial. El cero de este espacio es la forma bilineal

$$f: V \times V \to \mathbf{R}$$
$$f(x, y) = 0 \qquad \forall (x, y) \in V \times V$$

y el inverso aditivo de una forma $f: V \times V \to \mathbf{R}$ es la forma

$$-f: V \times V \to \mathbf{R}$$
$$(-f)(x, y) = -f(x, y)$$

Si se fija una base β del espacio V, se puede establecer por medio de ella una correspondencia entre los elementos del espacio vectorial $L_2(V, \mathbf{R})$ y las matrices cuadradas de orden n (los elementos del espacio vectorial $M_{n\times n}$), de la siguiente manera:

$$\varphi_\beta: L_2(V, \mathbf{R}) \to M_{n\times n}$$
$$\varphi_\beta(f) = [f]_\beta$$

El resultado principal de esta subsección establece que esta correspondencia es de hecho un **isomorfismo** de espacios vectoriales. Éste es precisamente el contenido del siguiente teorema.

TEOREMA 1.2

Sea B una base del espacio vectorial V de dimensión n. La función

$$\varphi_\beta: L_2(V, \mathbf{R}) \to M_{n\times n}$$
$$\varphi_\beta(f) = [f]_\beta$$

es un isomorfismo del espacio vectorial $L_2(V, \mathbf{R})$ al espacio vectorial $M_{n\times n}$.

DEMOSTRACIÓN φ_β es lineal: para $f_1, f_2 \in L_2(V, \mathbf{R})$ se tiene

$$\varphi_\beta(f_1 + f_2) = [f_1 + f_2]_\beta = [f_1]_\beta + [f_2]_\beta$$

en donde la última igualdad se justifica por la manera como se definió $f_1 + f_2$ y por la suma de matrices. Entonces,

$$\varphi_\beta(f_1 + f_2) = \varphi_\beta(f_1) + \varphi_\beta(f_2)$$

Similarmente, si $f \in L_2(V, \mathbf{R})$ y $c \in \mathbf{R}$ se tiene

$$\varphi_\beta(cf) = [cf]_\beta = c[f]_\beta = c\varphi_\beta(f)$$

lo que muestra que φ_β es lineal.

φ_β es sobreyectiva: dada $A \in M_{n\times n}$ defínase $f: V \times V \to \mathbf{R}$ como

$$f(x, y) = [x]_\beta^t A [y]_\beta$$

Al usar el hecho de que la función $w \mapsto [w]_\beta$ es un isomorfismo de V a $M_{n\times 1}$ es fácil ver que f así definida, es una forma bilineal en V.

Por último, φ_β es inyectiva: supóngase que para $f, g \in L_2(V, \mathbf{R})$ se tiene

$$[f]_\beta = [g]_\beta$$

Entonces, por la definición de matriz asociada a una forma bilineal en V respecto de la base $\beta = \{v_1, v_2, \ldots, v_n\}$ se concluye que

$$f(v_i, v_j) = g(v_i, v_j), \quad i, j = 1, 2, \ldots, n$$

Sean x, y dos vectores arbitrarios de V. Al escribir

$$x = \sum_{i=1}^{n} x_i v_i \qquad y = \sum_{j=1}^{n} y_j v_j$$

se tiene que

$$\begin{aligned}f(x, y) &= f\left(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j\right) = \sum_{i,j=1}^{n} f(v_i, v_j) x_i y_j = \sum_{i,j=1}^{n} g(v_i, v_j) x_i y_j\\ &= g\left(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j\right) = g(x, y)\end{aligned}$$

lo que muestra que $f = g$ y por lo tanto, que φ_β es inyectiva.

En resumen, φ_β es una transformación lineal biyectiva. Es decir, φ_β es un isomorfismo.

Q.E.D.

COROLARIO

> Si V es un espacio vectorial de dimensión n, se tiene que $dim\ L_2(V, \mathbf{R}) = n^2$.

DEMOSTRACIÓN Se sigue de inmediato del hecho de que $dim\ M_{n \times n} = n^2$ y del teorema anterior.

Q.E.D.

EJERCICIOS (SECCIÓN 1, CAPÍTULO 7)

1. Determine cuáles de las siguientes funciones f: $\mathbf{R}^2 \to \mathbf{R}$ son formas bilineales ($x = (x_1, x_2)$, $y = (y_1, y_2)$).

   a) $f(x, y) = x_1 + y_1$
   b) $f(x, y) = x_1y_1 + x_2y_2$
   c) $f(x, y) = x_1y_2 - y_1x_2$
   d) $f(x, y) = (x_1 + y_1)^2 - (x_1 - y_1)^2$
   e) $f(x, y) = 5$

2. Sea V el espacio vectorial de las matrices de orden 2×3. Para $A, B \in V$, defina f: $V \times V \to \mathbf{R}$ como

$$f(A, B) = tr(A^tB)$$

   Demuestre que f es una forma bilineal en V. Más generalmente, si C es una matriz fija de orden 2, demuestre que la función f: $V \times V \to \mathbf{R}$

$$f(A, B) = tr(A^tCB)$$

   es una forma bilineal en V.

3. Sea V un espacio vectorial con producto interno y sean u_1 y u_2 dos vectores fijos de V, compruebe que la función f: $V \times V \to \mathbf{R}$ dada por

$$f(v_1, v_2) = (v_1 \mid u_1)(v_2 \mid u_2)$$

   es una forma bilineal en V.

4. Sea $V = M_{n \times n}$, demuestre que la función f: $V \times V \to \mathbf{R}$ dada por

$$f(A, B) = n\ tr(AB) - (tr\ A)(tr\ B)$$

   es una forma bilineal en V.

5. Escriba cada una de las siguientes formas bilineales en $\mathbf{R}^3$, f: $\mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$ como $f(x, y) = [x]^t_\beta A[y]_\beta$, en donde A es la matriz de la forma f respecto de la base β (β es la base

canónica de $\mathbf{R}^3$).

   a) $f(x, y) = x_1y_1$
   b) $f(x, y) = x_1y_1 + x_2y_2$
   c) $f(x, y) = x_1y_1 + x_2y_2 + x_3y_3$
   d) $f(x, y) = x_1y_1 + x_1y_2 + x_1y_3 + x_2y_1 + x_2y_2 + x_2y_3 + x_3y_1 + x_3y_2 + x_3y_3$
   e) $f(x, y) = 3x_1y_3 + 2x_2y_2 + x_3y_1$
   f) $f(x, y) = 3x_1y_1 + 8x_1y_2 - 7x_1y_3 + 10x_2y_3 - 12x_2y_2 + 9x_2y_1 + x_3y_2 + 7x_3y_1 - 10x_3y_3$

6. Sea V el espacio vectorial $\mathbf{R}^3$ con el producto interno canónico. Sean $u_1 = (1, 0, 1)$, $u_2 = (0, 2, -1)$ dos vectores de $\mathbf{R}^3$. Considere la forma bilineal f: $\mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$ del ejercicio 3

$$f(v_1, v_2) = (v_1 \mid u_1)(v_2 \mid u_2)$$

   a) Hallar la matriz de la forma f respecto de la base canónica de $\mathbf{R}^3$.
   b) Encontrar la matriz de la forma f respecto de la base β de $\mathbf{R}^3$ dada por

$$\beta = \{(3, 0, 0), (5, 2, 1), (-1, 1, 1)\}$$

7. Sea $V = M_{2 \times 2}$ y considere la forma bilineal f: $V \times V \to \mathbf{R}$ del ejercicio 4

$$f(A, B) = 2\ tr(AB) - (tr\ A)(tr\ B)$$

   a) Hallar la matriz de la forma f respecto de la base canónica de $M_{2 \times 2}$
   b) Encontrar la matriz de la forma f respecto de la base β de $M_{2 \times 2}$ dada por

$$\beta = \left\{\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 0 & 3 \end{bmatrix}\right\}$$

8. Considere la forma bilineal f del ejercicio 2, f: $M_{2 \times 3} \times M_{2 \times 3} \to \mathbf{R}$

$$f(A, B) = tr(A^tCB)$$

   en donde

$$C = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}$$

   a) Hallar la matriz de la forma f respecto de la base canónica de $M_{2 \times 3}$.
   b) Encontrar la matriz de la forma f respecto de la base β de $M_{2 \times 3}$ dada por

$$\beta = \left\{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\right\}$$

9. Sea V un espacio vectorial de dimensión 4 y sea f: $V \times V \to \mathbf{R}$ una forma bilineal en V que respecto de la base $\beta = \{v_1, v_2, v_3, v_4\}$ de V, tiene por matriz asociada a

$$A = \begin{bmatrix} 1 & -1 & 2 & 3 \\ 4 & 0 & 2 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

Halle la matriz de la forma f respecto de la base de V dada por

a) $\beta_1 = \{v_2, v_4, v_3, v_1\}$

b) $\beta_2 = \{v_1, v_1 + v_2, v_1 + v_2 + v_3, v_1 + v_2 + v_3 + v_4\}$

c) $\beta_3 = \{v_1 + v_2 + v_3 + v_4, v_1 - v_2 + v_3 + v_4, v_1 + v_2 - v_3 + v_4, v_1 + v_2 + v_3 - v_4\}$

10. Sea $V = \mathbf{R}^3$ y considere la forma bilineal f: $\mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$ dada por

$$f(x, y) = x_1y_2 + x_1y_3 + x_2y_1 + x_2y_3 + x_3y_1 + x_3y_2 \qquad \begin{pmatrix} x = (x_1, x_2, x_3) \\ y = (y_1, x_2, y_3) \end{pmatrix}$$

a) Encuentre la matriz asociada a la forma f respecto de la base canónica de $\mathbf{R}^3$.

b) Halle la matriz asociada a la forma f respecto de la base β de $\mathbf{R}^3$

$$\beta = \{(2/3, 1/3, -2/3), (2/3, -2/3, 1/3), (1/3, 2/3, 2/3)\}$$

c) Demuestre que las matrices de los dos incisos anteriores tienen la misma traza y el mismo determinante.

③ 11. Sea V un espacio vectorial de dimensión finita y sean β y β' bases de V. Considere la forma bilineal f: $V \times V \to \mathbf{R}$. ¿Cierto o falso?: la matriz asociada a la forma f respecto de la base β tiene la misma traza y el mismo determinante de la matriz asociada a f respecto de la base β'. Concilie su respuesta con el resultado del ejercicio anterior.

[*Sugerencia*: las matrices de los incisos (a) y (b) en el ejercicio 10 son *semejantes*. ¿Es esto cierto en general para una forma bilineal?]

12. Sea V un espacio vectorial de dimensión n y sea β una base cualquiera de V. Considere la forma bilineal f: $V \times V \to \mathbf{R}$ y suponga que se tiene la propiedad

$$f(v_1, v_2) = f(v_2, v_1) \qquad \forall\, v_1, v_2 \in V$$

Sea A la matriz asociada a la forma f respecto de la base β. Pruebe que A es una matriz simétrica. ¿Cómo se altera esta propiedad si se cambia la base de V?

13. ¿Cuáles de las formas bilineales del ejercicio 5 son no degeneradas?

14. Determine el rango de la forma bilineal f: $\mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$ del ejercicio 6

a) Usando la matriz de la forma f obtenida en el inciso (a).

b) Usando la matriz de la forma f obtenida en el inciso (b).

15. Demuestre que la forma bilineal del ejercicio 7 es degenerada.

16. Determine el rango de la forma bilineal del ejercicio 8.

17. Determine el rango de la forma bilineal del ejercicio 9.

18. Sea V un espacio vectorial de dimensión 3. Considere la forma bilineal en V, f: $V \times V \to \mathbf{R}$. Suponga que respecto de alguna base β de V, esta forma tiene asociada la matriz

$$A = \begin{bmatrix} 4 & -3 & 5 \\ 1 & -7 & -5 \\ 3 & -2 & 8 \end{bmatrix}$$

Compruebe que no existe base β' de V, respecto de la cual la matriz asociada a la forma bilineal f sea

$$B = \begin{bmatrix} 3 & 4 & -5 \\ 8 & 7 & -2 \\ 2 & -1 & 8 \end{bmatrix}$$

(*Sugerencia*: ¿Cuál es el rango de f?)

③ 19. Sea V un espacio vectorial de dimensión finita y sea f: $V \times V \to \mathbf{R}$ una forma bilineal en V. Demuestre que el subconjunto W de V dado por

$$W = \{x \in V \mid f(x, y) = 0 \qquad \forall\, y \in V\}$$

es un subespacio de V. Más aún, demuestre que rango de $f = \dim V - \dim W$. Concluya entonces que la forma bilineal f es no degenerada si, y sólo si W es el subespacio trivial de V.

20. Describa el subespacio W del ejercicio anterior para la forma bilineal, f: $V \times V \to \mathbf{R}$, $f(x, y) = 0$, $\forall\, x, y \in V$.

21. Describa el subespacio W del ejercicio 19 para cada una de las siguientes formas bilineales f: $\mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$. Calcule en cada caso el rango de la forma usando la fórmula establecida en el ejercicio 19.

a) $f(x, y) = x_1y_1$

b) $f(x, y) = 2x_1y_1 + 3x_2y_2$

c) $f(x, y) = 6x_1y_1 - 4x_2y_2 + 10x_3y_3$

d) $f(x, y) = 7x_1y_1 + 8x_2y_3 - 6x_3y_2$

e) $f(x, y) = x_1y_1 + x_1y_2 + x_1y_3 + 2x_2y_1 + 2x_2y_2 + 2x_2y_3 + 7x_3y_1 + 9x_3y_2 + 10x_3y_3$

22. Considere la forma bilineal del ejercicio 3. Suponga que los vectores u_1 y u_2 son vectores (fijos) no nulos de V. Describa el subespacio W definido en el ejercicio 16.

23. Considere las formas bilineales en $\mathbf{R}^3$

$$f_1, f_2: \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R} \qquad \begin{aligned} f_1(x, y) &= x_1y_1 + 3x_2y_2 - 2x_3y_3 + 4x_2y_1 \\ f_2(x, y) &= 6x_1y_1 + 7x_1y_3 + 8x_2y_1 + 6x_2y_3 - x_3y_1 + x_3y_2 \end{aligned}$$

a) Determine una fórmula explícita para la forma

$$f_1 + 3f_2: \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$$

b) Si β es la base canónica de $\mathbf{R}^3$, verifique que

$$[f_1 + 3f_2]_\beta = [f_1]_\beta + 3[f_2]_\beta$$

④ 24. (*Un ejercicio sobre aplicaciones bilineales.*) Sean V_1, V_2 y U tres espacios vectoriales. Se dice que la función f: $V_1 \times V_2 \to U$ es una *aplicación bilineal* si es lineal respecto de cada una de sus dos variables, cuando la otra se mantiene fija. Más precisamente, f: $V_1 \times V_2 \to U$ es una aplicación bilineal si

$$f(v_1 + cv_1', v_2) = f(v_1, v_2) + cf(v_1', v_2)$$

y

$$f(v_1, v_2 + cv_2') = f(v_1, v_2) + cf(v_1, v_2')$$

para todo $v_1, v_1' \in V_1$, $v_2, v_2' \in V_2$ y $c \in \mathbf{R}$.

a) Sea V un espacio vectorial y sea f: $\mathbb{R} \times V \to V$ la función dada por $f(c, v) = cv$. Demuestre que f es una aplicación bilineal.

b) Sean V y U dos espacios vectoriales y T: $V \to U$ una transformación lineal. Considere la función f: $L(V, U) \times V \to U$ (en donde $L(V, U)$ es el espacio vectorial de las transformaciones lineales de V a U) dada por $f(T, v) = T(v)$. Demuestre que f es una aplicación bilineal.

c) Sean V_1, V_2, U_1 y U_2 espacios vectoriales y T: $U_1 \to U_2$ una transformación lineal. Suponga que f: $V_1 \times V_2 \to U_1$ es una aplicación bilineal. Demuestre entonces que la composición $T \circ f$: $V_1 \times V_2 \to U_2$ es una aplicación bilineal.

d) Sean V_1, V_2, V_3 V_4 y U espacios vectoriales T_1: $V_1 \to V_3$ y T_2: $V_2 \to V_4$ transformaciones lineales. Suponga que f: $V_3 \times V_4 \to U$ es una aplicación bilineal. Defina g: $V_1 \times V_2 \to U$ como $g(v_1, v_2) = f(T(v_1), T(v_2))$. Demuestre que g es una aplicación bilineal.

e) Defina la *imagen* de una aplicación bilineal F: $V_1 \times V_2 \to U$, denotada por $Im\ F$ como el conjunto

$$Im\ F = \{u \in U \mid \exists\ v_1 \in V_1, v_2 \in V_2, \text{ tales que } F(v_1, v_2) = u\}$$

En este inciso se verá que este subconjunto de U no es en general un subespacio de U, al contrario de lo que acontece con las aplicaciones (transformaciones) lineales T: $V \to U$ cuya imagen siempre es un subespacio de su codominio (del espacio vectorial U).

e1) Considere la función F: $(\mathbb{R}^2)^* \times (\mathbb{R}^2)^* \to L_2(\mathbb{R}^2, \mathbb{R})$ (en donde $(\mathbb{R}^2)^*$ denota el espacio dual de $\mathbb{R}^2$ y $L_2(\mathbb{R}^2, \mathbb{R})$ indica el espacio de formas bilineales en $\mathbb{R}^2$ estudiado en la subsección 1.3) dada por

$$F(f, g)(x, y) = f(x)g(y) \qquad f, g \in (\mathbb{R}^2)^*, x, y \in \mathbb{R}^2$$

Verifique primeramente que $F(f, g)$ es un elemento de $L_2(\mathbb{R}^2, \mathbb{R})$. Compruebe luego que F es una aplicación bilineal.

e2) Sea $\{e_1, e_2\}$ la base canónica de $\mathbb{R}^2$ y $\{f_1, f_2\}$ su base dual. Observe que $F(f_1, f_1)$ y $F(f_2, f_2)$ pertenecen (obviamente) a la imagen de F. Se verá que $F(f_1, f_1) + F(f_2, f_2)$ no pertenece a la imagen de F, mostrando así que $Im\ F$ no es un subespacio de U. Suponga, para llegar a una contradicción, que $F(f_1, f_1) + F(f_2, f_2)$ sí pertenece a $Im\ F$. En tal caso, existirían $f, g \in (\mathbb{R}^2)^*$ tales que

$$F(f, g) = F(f_1, f_1) + F(f_2, f_2)$$

Evalúe esta expresión en (e_1, e_1) y en (e_2, e_2) para concluir que los números reales $f(e_1), f(e_2), g(e_1)$ y $g(e_2)$ deben ser distintos de cero.

e3) Evalúe ahora la misma expresión en (e_1, e_2) para obtener la contradicción deseada.

f) Denote por $L(V_1, V_2; U)$ al conjunto de todas las aplicaciones bilineales del tipo f: $V_1 \times V_2 \to U$. Sean $f_1, f_2 \in L(V_1, V_2; U)$ y $c \in \mathbb{R}$. Defina $f_1 + f_2$: $V_1 \times V_2 \to U$ y cf_1: $V_1 \times V_2 \to U$ como

$$(f_1 + f_2)(v_1, v_2) = f_1(v_1, v_2) + f_2(v_1, v_2)$$
$$(cf_1)(v_1, v_2) = cf_1(v_1, v_2)$$

Verifique que tanto $f_1 + f_2$ como cf_1 pertenecen a $L(V_1, V_2; U)$ y que, de hecho, el conjunto $L(V_1, V_2; U)$ con estas operaciones es un espacio vectorial.

g) Considere la función T: $L(V_1, V_2; U) \to L(V_1, L(V_2, U))$ [en donde se está denotando por $L(W, W')$ al espacio de transformaciones lineales de W a W'] dada por

$$((Tf)(v_1))(v_2) = f(v_1, v_2) \qquad v_1 \in V_1, v_2 \in V_2$$

g1) Demuestre que T es una transformación lineal.

g2) Demuestre que T es inversible, mostrando que T^{-1}: $L(V_1, L(V_2, U)) \to L(V_1, V_2; U)$

$$T^{-1}(g)(v_1, v_2) = (g(v_1))(v_2)$$

es la transformación inversa de T. Concluya entonces que T es un isomorfismo.

g3) Concluya que si V_1, V_2 y U son espacios vectoriales de dimensión finita, entonces

$$dim\ L(V_1, V_2; U) = (dim\ V_1)(dim\ V_2)(dim\ U)$$

Use este hecho para obtener nuevamente el resultado del corolario del teorema 1.2 de esta sección.

2. FORMAS CUADRÁTICAS

En la sección anterior se presentó el concepto de forma bilineal en un espacio vectorial V. Se analizó la relación que existe entre las formas bilineales en un espacio V de dimensión finita y el espacio de matrices cuadradas de orden $dim\ V$. En esta sección se introdujeron las formas cuadráticas como un tipo especial de formas bilineales. Sobre ellas se trabajará en el resto del capítulo.

DEFINICIÓN 2.1 Se dice que la forma bilineal f: $V \times V \to \mathbb{R}$ en el espacio vectorial V es una *forma bilineal simétrica* si

$$f(x, y) = f(y, x) \qquad \forall\ x, y \in V$$

Si V es de dimensión finita y β es una base de V, es fácil ver que la matriz asociada a una forma bilineal simétrica f en V es una matriz simétrica. En efecto, por definición se tiene que el i, j-ésimo elemento de tal matriz es $f(v_i, v_j)$ en donde $\beta = \{v_1, v_2, \ldots, v_n\}$ es la base de V. Pero siendo f simétrica se tiene que $f(v_i, v_j) = f(v_j, v_i)\ \forall\ i, j = 1, 2, \ldots, n$, lo que muestra que la matriz asociada a f respecto de la base β es simétrica.

Recíprocamente, si la matriz asociada a una forma bilineal f: $V \times V \to \mathbb{R}$ en el espacio vectorial V de dimensión finita respecto de una base β de V es una matriz simétrica, la forma f es una forma bilineal simétrica, pues siendo una matriz de orden 1 simétrica se puede escribir

$$f(x, y) = [x]^t_\beta A[y]_\beta = ([x]^t_\beta A[y]_\beta)^t = [y]^t_\beta A^t[x]_\beta = [y]^t_\beta A[x]_\beta = f(y, x)$$

DEFINICIÓN 2.2 Una *forma cuadrática* en el espacio vectorial V es una función $q: V \to \mathbf{R}$ dada por

$$q(v) = f(v, v) \tag{2.1}$$

en donde f es una forma bilineal simétrica en V.

Por ejemplo, si $V = \mathbf{R}^2$, considérese la forma bilineal $f: \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ dada por

$$f(x, y) = [x_1\, x_2] \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Como la matriz

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$

es simétrica, la forma f también lo es. La función $q: \mathbf{R}^2 \to \mathbf{R}^2$ dada por

$$q(x_1, x_2) = f(x, x) = [x_1\, x_2] \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 6x_1x_2 + 5x_2^2$$

es entonces una forma cuadrática en $\mathbf{R}^2$.

Se tiene, pues, que las formas cuadráticas son generadas por formas bilineales simétricas, en el sentido de que a cada forma bilineal simétrica f le corresponde una forma cuadrática (según la definición 2.1) a la que se llama "forma cuadrática generada por la forma bilineal simétrica f".

Es importante observar que una forma cuadrática determina *de modo único* a la forma bilineal simétrica que la genera. En efecto, supóngase que $q: V \to \mathbf{R}$ es una forma cuadrática en V (no necesariamente de dimensión finita). Se afirma que la única forma bilineal simétrica $f: V \times V \to \mathbf{R}$ que genera a q es:

$$f(x, y) = \frac{1}{4}q(x + y) - \frac{1}{4}q(x - y) \tag{2.2}$$

Para ver esto, considérese la forma bilineal simétrica $\tilde{f}: V \times V \to \mathbf{R}$ tal que $q(x) = \tilde{f}(x, x)$. Entonces,

$$q(x + y) = \tilde{f}(x + y, x + y) = \tilde{f}(x, x) + 2\tilde{f}(x, y) + \tilde{f}(y, y)$$
$$q(x - y) = \tilde{f}(x - y, x - y) = \tilde{f}(x, x) - 2\tilde{f}(x, y) + \tilde{f}(y, y)$$

Al restar estas expresiones y despejar $\tilde{f}(x, y)$ se obtiene

$$\tilde{f}(x, y) = \frac{1}{4}q(x + y) - \frac{1}{4}q(x - y) = f(x, y)$$

lo que muestra que $f = \tilde{f}$.

A la fórmula (2.2) se le conoce como *identidad de polarización*. Debido a ella se puede entonces hablar de *la* forma bilineal simétrica que genera a una forma cuadrática q.

Según el análisis anterior, una vez fijada una base del espacio vectorial V de dimensión finita, se puede hacer una identificación biyectiva entre formas cuadráticas en V y matrices cuadradas simétricas de orden $dim\ V$ (¿por qué?).

Si V es un espacio vectorial de dimensión n y $q: V \to \mathbf{R}$ es una forma cuadrática en V, se puede entonces escribir

$$q(x) = [x]_\beta^t A[x]_\beta \tag{2.3}$$

en donde β es una base de V y A es una matriz simétrica bien determinada por q (y por la base β, claro está) con elementos a_{ij}.

Si

$$[x]_\beta = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

se tiene entonces que

$$q(x) = [x_1\, x_2 \ldots x_n] \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12} & a_{22} & \ldots & a_{2n} \\ & & & \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i,j=1}^{n} a_{ij}x_ix_j$$

$$= \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j}^{n} a_{ij}x_ix_j \tag{2.4}$$

Ésta es, pues, la estructura algebraica que posee una forma cuadrática en un espacio vectorial de dimensión finita.

De ella se puede ver la relación que existe entre los coeficientes de la expresión algebraica de la forma y los elementos de la matriz simétrica que le corresponde.

EJEMPLO 1 Por ejemplo, la forma cuadrática $q: \mathbf{R}^3 \to \mathbf{R}$ dada por

$$q(x_1, x_2, x_3) = x_1^2 + 3x_2^2 - x_3^2 - 8x_1x_2 + 4x_1x_3 + 2x_2x_3$$

se puede escribir como

$$q(x_1, x_2, x_3) = x_1^2 + 3x_2^2 - x_3^2 + 2(-4x_1x_2 + 2x_1x_3 + x_2x_3)$$

y así, al comparar con (2.4), a esta forma le corresponde, respecto de la base canónica de $\mathbf{R}^3$ la matriz simétrica 3×3

$$A = \begin{bmatrix} 1 & -4 & 2 \\ -4 & 3 & 1 \\ 2 & 1 & -1 \end{bmatrix}$$

y entonces se puede escribir

$$q(x) = X^t A X$$

Se puede ver fácilmente que la forma bilineal simétrica que genera a esta forma cuadrática es:

$$f(x, y) = X^t A Y = x_1y_1 - 4x_1y_2 + 2x_1y_3 - 4x_2y_1 + 3x_2y_2 + x_2y_3 + 2x_3y_1 + x_3y_2 - x_3y_3$$

2.1. REDUCCIÓN A UNA SUMA DE CUADRADOS

Considérese una forma cuadrática $q: V \to \mathbf{R}$ en el espacio vectorial V de dimensión n. Si β es una base de V, se puede escribir para la matriz simétrica A determinada por q

$$q(x) = [x]_\beta^t A[x]_\beta$$

Si se toma una nueva base β' de V, la matriz de la forma bilineal correspondiente (la que genera a q) respecto de esta nueva base sería P^tAP en donde P es la matriz de cambio de base de β' a β. En tal caso se tendría entonces,

$$q(x) = [x]_{\beta'}^t P^tAP[x]_{\beta'}$$

EJEMPLO 2 Por ejemplo, sea $q: \mathbf{R}^2 \to \mathbf{R}$ la forma cuadrática

$$q(x_1, x_2) = x_1^2 + x_2^2 + 4x_1x_2$$

Respecto de la base canónica β de $\mathbf{R}^2$ la matriz asociada a q es:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

si se considera ahora la base $\mathbf{R}^2$

$$\beta' = \{(1/\sqrt{2}, -1/\sqrt{2}), (1/\sqrt{2}, 1/\sqrt{2})\}$$

se tiene que la matriz de cambio de base de β' a β es:

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

de modo que la matriz asociada a la forma q respecto de la base β' es:

$$P^tAP = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$$

y así, la expresión para la forma cuadrática q en esta nueva base es:

$$q(x) = [x]_{\beta'}^t P^tAP[x]_{\beta'} = [x_1' \; x_2'] \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \end{bmatrix}$$
$$= -(x_1')^2 + 3(x_2')^2$$

en donde $(x)_{\beta'} = (x_1', x_2')$.

Observe que respecto de la base β', la forma cuadrática q se puede expresar como una suma de cuadrados (de las coordenadas del vector x en que se evalúa q, respecto de la base β').

El siguiente teorema establece que siempre es posible expresar una forma cuadrática $q: V \to \mathbf{R}$ en el espacio vectorial V de dimensión finita como una suma de cuadrados.

TEOREMA 2.1 Sea $q: V \to \mathbf{R}$ una forma cuadrática en el espacio vectorial V de dimensión n. Existe una base β de V tal que

$$q(x) = \sum_{i=1}^{n} \lambda_i x_i^2$$

en donde $(x)_\beta = (x_1, x_2, \ldots, x_n)$.

DEMOSTRACIÓN Sea A la matriz asociada a la forma q respecto de alguna (cualquiera) base $\tilde{\beta}$ de V. Es decir, que q puede escribirse como

$$q(x) = [x]_{\tilde{\beta}}^t A[x]_{\tilde{\beta}}$$

Siendo A una matriz simétrica, se sabe que existe una matriz ortogonal P tal que P^tAP es una matriz diagonal (teorema 4.3 del capítulo 6); dígase que

$$P^tAP = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

Sea β la base de V tal que P es la matriz de cambio de base de β a $\tilde{\beta}$.

La expresión de la forma q respecto de esta base β es entonces,

$$q(x) = [x]^t_{\beta} P^t AP[x]_{\beta} = [x_1\, x_2 \ldots x_n] \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2$$

en donde $(x)_{\beta} = (x_1, x_2, \ldots, x_n)$. Esto es precisamente lo que se quería demostrar.

Q.E.D.

COROLARIO

(DE LA DEMOSTRACIÓN DEL TEOREMA ANTERIOR.) Sea $q: V \to \mathbf{R}$ una forma cuadrática en el espacio vectorial V de dimensión n. Sea A la matriz de la forma q respecto de alguna base $\tilde{\beta}$ de V. Entonces existe una base β de V para la cual

$$q(x) = \sum_{i=1}^{n} \lambda_i x_i^2 \qquad (2.5)$$

en donde $(x)_{\beta} = (x_1, x_2, \ldots, x_n)$ y $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los valores propios de la matriz A.

DEMOSTRACIÓN Al tomar la base β de V, como en la demostración del teorema anterior, tal que la matriz ortogonal P es la matriz de cambio de base de β a $\tilde{\beta}$, se observa que P diagonaliza ortogonalmente a la matriz A. Es decir, que los elementos de la diagonal principal de la matriz P^tAP son los valores propios de A.

Q.E.D.

EJEMPLO 3 Considere por ejemplo, la forma cuadrática $q: \mathbf{R}^3 \to \mathbf{R}$

$$q(x_1, x_2, x_3) = 3x_2^2 + 3x_3^2 + 4x_1x_2 + 4x_1x_3 - 2x_2x_3 \qquad (2.6)$$

Se puede escribir la expresión para q como

$$q(x_1, x_2, x_3) = [x_1\, x_2\, x_3] \begin{bmatrix} 0 & 2 & 2 \\ 2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

de modo que la matriz asociada a la forma q respecto de la base canónica de $\mathbf{R}^3$ es:

$$A = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$$

Diagonalice ortogonalmente a la matriz A. El polinomio característico de esta matriz es:

$$p(\lambda) = det(A - \lambda I) = det \begin{bmatrix} -\lambda & 2 & 2 \\ 2 & 3-\lambda & -1 \\ 2 & -1 & 3-\lambda \end{bmatrix} = -\lambda^3 + 6\lambda^2 - 32$$

$$= -(\lambda - 4)^2(\lambda + 2)$$

de modo que los valores propios de A son $\lambda_1 = \lambda_2 = 4$ y $\lambda_3 = -2$.

Para $\lambda = 4$ se obtienen los vectores propios

$$v_1 = (1, 2, 0)$$

$$v_2 = (1, 0, 2)$$

Al aplicar el proceso de Gram-Schmidt a estos vectores se obtienen los vectores ortonormales*

$$u_1 = \frac{1}{\sqrt{5}}(1, 2, 0)$$

$$u_2 = \frac{1}{\sqrt{30}}(2, -1, 5)$$

Para el valor propio $\lambda = -2$ se obtiene el vector propio

$$v_3 = (-2, 1, 1)$$

que normalizado queda

$$u_3 = \frac{1}{\sqrt{6}}(-2, 1, 1)$$

Entonces, la matriz que diagonaliza ortogonalmente a la matriz A es:

$$P = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{30} & -2/\sqrt{6} \\ 2/\sqrt{5} & -1/\sqrt{30} & 1/\sqrt{6} \\ 0 & 5/\sqrt{30} & 1/\sqrt{6} \end{bmatrix}$$

Si se considera entonces, la nueva base de $\mathbf{R}^3$

$$\beta = \{u_1, u_2, u_3\} = \left\{ \frac{1}{\sqrt{5}}(1, 2, 0), \frac{1}{\sqrt{30}}(2, -1, 5), \frac{1}{\sqrt{6}}(-2, 1, -1) \right\}$$

*Se está considerando el producto interno canónico de $\mathbf{R}^3$.

se tiene que la forma q se presenta en esta base como

$$q(x) = [x]_\beta^t P^t A P [x]_\beta = [x_1' \; x_2' \; x_3'] \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix}$$

$$= 4(x_1')^2 + 4(x_2')^2 - 2(x_3')^2 \qquad (2.7)$$

en donde $(x)_\beta = (x_1', x_2', x_3')$. Es decir, respecto de la base de β la forma cuadrática q es una suma de cuadrados.

Observe que las coordenadas x_1', x_2' y x_3' de x respecto de la base β están relacionadas con las coordenadas x_1, x_2, x_3 de x respecto de la base canónica de $\mathbf{R}^3$ como*

$$x_1' = \frac{x_1 + 2x_2}{\sqrt{5}}$$

$$x_2' = \frac{2x_1 - x_2 + 5x_3}{\sqrt{30}} \qquad (2.8)$$

$$x_3' = \frac{-2x_1 + x_2 + x_3}{\sqrt{6}}$$

Desde otro punto de vista, se puede contemplar el ejemplo anterior como un "proceso de cambio de variables en la forma cuadrática q": dada la expresión (2.6) de la forma q, se ha realizado en ella el cambio de las variables x_1, x_2, x_3 por las variables x_1', x_2', x_3' según las expresiones (2.8) de tal modo, que q queda así escrita como una suma de cuadrados, según la expresión (2.7).

Se debe hacer la aclaración de que la base β de la que habla el teorema 2.1, respecto de la cual la forma cuadrática q queda escrita como una suma de cuadrados *no es única*. Con la perspectiva del comentario del párrafo anterior, esto significa que existen varias maneras de realizar cambios de variables en la expresión algebraica de una forma cuadrática de modo que ésta queda escrita como una suma de cuadrados (en las nuevas variables). Sin embargo, es claro que sí es única la base β establecida en el corolario del teorema 2.1, respecto de la cual la forma q queda escrita como una suma de cuadrados siendo los coeficientes de éstos los valores propios de la matriz asociada a q respecto de alguna base dada de V.

Retome el ejemplo presentado anteriormente. Se tiene la forma: $q: \mathbf{R}^3 \to \mathbf{R}$

$$q(x_1, x_2, x_3) = 3x_2^2 + 3x_3^2 + 4x_1x_2 + 4x_1x_3 - 2x_2x_3$$

$$= [x_1 \; x_2 \; x_3] \begin{bmatrix} 0 & 2 & 2 \\ 2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

*Siendo β una base ortonormal de $\mathbf{R}^3$ con el producto interno canónico, se tiene que si $x = (x_1, x_2, x_3)$ entonces $x' = (x \mid u_i)$, $i = 1, 2, 3$.

Si se considera la base β de $\mathbf{R}^3$ dada por

$$\beta = \left\{(1, 1, 0), \left(-\frac{5}{7}, \frac{2}{7}, 0\right), (2, -1, 1)\right\}$$

se tiene que la matriz de cambio de base de β a la base canónica de $\mathbf{R}^3$ es:

$$P = \begin{bmatrix} 1 & -5/7 & 2 \\ 1 & 2/7 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

de modo que la matriz asociada a la forma q respecto de la nueva base es:

$$P^t A P = \begin{bmatrix} 1 & 1 & 0 \\ -5/7 & 2/7 & 0 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 2 & 2 \\ 2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & -5/7 & 2 \\ 1 & 2/7 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 0 & 0 \\ 0 & -4/7 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

y entonces,

$$q(x) = [x]_\beta^t P^t A P [x]_\beta = [x_1' \; x_2' \; x_3'] \begin{bmatrix} 7 & 0 & 0 \\ 0 & -4/7 & 0 \\ 0 & 0 & 8 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix}$$

$$= 7(x_1')^2 - \frac{4}{7}(x_2')^2 + 8(x_3')^2 \qquad (2.9)$$

en donde $(x)_\beta = (x_1', x_2', x_3')$. Al hacer algunos cálculos sencillos se puede llegar a las expresiones de "cambio de variables"

$$x_1' = \frac{2}{7}x_1 + \frac{5}{7}x_2 + \frac{1}{7}x_3$$

$$x_2' = -x_1 + x_2 + 3x_3 \qquad (2.10)$$

$$x_3' = x_3$$

Entonces, respecto de esta nueva base β de $\mathbf{R}^3$, la forma q también se presenta como una suma de cuadrados.

Termina esta subsección indicando *cómo* fue posible establecer la fórmula (2.9) y la base β anterior a partir de la expresión (2.6) de la forma cuadrática q:*

Como la expresión (2.6) de la forma q (la cual se quiere escribir como una suma de cuadrados) no posee el término correspondiente al cuadrado de la primera coordenada, se hace primeramente el cambio de variable.

$$x_1 = y_1$$

*No se pretende dar la versión general de este procedimiento. Sin embargo, el lector podrá sin mucha dificultad aplicar lo aquí mostrado en este ejemplo concreto a otro tipo de ejemplos.

$$x_2 = y_1 + y_2$$
$$x_3 = y_3$$

quedando entonces,

$$\begin{aligned} q(x) &= 3x_2^2 + 3x_3^2 + 4x_1x_2 + 4x_1x_3 - 2x_2x_3 \\ &= 3(y_1+y_2)^2 + 3y_3^2 + 4y_1(y_1+y_2) + 4y_1y_3 - 2(y_1+y_2)y_3 \\ &= 7y_1^2 + 3y_2^2 + 3y_3^2 + 10y_1y_2 + 2y_1y_3 - 2y_2y_3 \end{aligned}$$

Se comienza ahora a "completar cuadrados" iniciando con y_1: se factoriza el coeficiente de y_1^2 y los términos que contengan y_1 quedando como

$$q(x) = 7\left(y_1^2 + \frac{2}{7}(5y_2 + y_3)y_1\right) + 3y_2^2 + 3y_3^2 - 2y_2y_3$$

se completa el cuadrado para y_1:

$$\begin{aligned} q(x) &= 7\left(y_1^2 + \frac{2}{7}(5y_2+y_3)y_1 + \frac{1}{49}(5y_2+y_3)^2\right) + 3y_2^2 + 3y_3^2 - 2y_2y_3 - \frac{1}{7}(5y_2+y_3)^2 \\ &= 7\left(y_1 + \frac{1}{7}(5y_2+y_3)\right)^2 + 3y_2^2 + 3y_3^2 - 2y_2y_3 - \frac{1}{7}(5y_2+y_3)^2 \\ &= 7\left(y_1 + \frac{5}{7}y_2 + \frac{1}{7}y_3\right)^2 - \frac{4}{7}y_2^2 + \frac{20}{7}y_3^2 - \frac{24}{7}y_2y_3 \end{aligned}$$

Se hace lo mismo con y_2:

$$\begin{aligned} q(x) &= 7\left(y_1 + \frac{5}{7}y_2 + \frac{1}{7}y_3\right)^2 - \frac{4}{7}(y_2^2 + 6y_2y_3 + 9y_3^2) + \frac{20}{7}y_3^2 + \frac{36}{7}y_3^2 \\ &= 7\left(y_1 + \frac{5}{7}y_2 + \frac{1}{7}y_3\right)^2 - \frac{4}{7}(y_2 + 3y_3)^2 + 8y_3^2 \end{aligned}$$

Al realizar

$$\begin{aligned} x_1' &= y_1 + \frac{5}{7}y_2 + \frac{1}{7}y_3 \\ x_2' &= y_2 + 3y_3 \\ x_3' &= y_3 \end{aligned}$$

o bien, en términos de x_1, x_2, x_3

$$\begin{aligned} x_1' &= \frac{2}{7}x_1 + \frac{5}{7}x_2 + \frac{1}{7}x_3 \\ x_2' &= -x_1 + x_2 + 3x_3 \\ x_3' &= x_3 \end{aligned}$$

se obtiene finalmente la expresión deseada (2.9) que presenta a q como una suma de cuadrados. Ésta es:

$$q(x) = 7(x_1')^2 - \frac{4}{7}(x_2')^2 + 8(x_3')^2$$

Para investigar cuál es la base β de $\mathbf{R}^3$ respecto de la cual q se presenta como en la expresión anterior se puede proceder como sigue: se sabe que $(x)_\beta = (x_1', x_2', x_3')$ en donde x_1', x_2', x_3' están dadas por las fórmulas (2.1). Sea $\beta = \{v_1, v_2, v_3\}$ la base procurada. Se debe tener entonces, que

$$\begin{aligned} (x_1, x_2, x_3) &= x_1'v_1 + x_2'v_2 + x_3'v_3 \\ &= \left(\frac{2}{7}x_1 + \frac{5}{7}x_2 + \frac{1}{7}x_3\right)v_1 + (-x_1 + x_2 + 3x_3)v_2 + x_3v_3 \end{aligned}$$

Al escribir los vectores de la base canónica de $\mathbf{R}^3$ en la expresión anterior se obtiene

$$\begin{aligned} (1,0,0) &= \frac{2}{7}v_1 - v_2 \\ (0,1,0) &= \frac{5}{7}v_1 + v_2 \\ (0,0,1) &= \frac{1}{7}v_1 + 3v_2 + v_3 \end{aligned}$$

de donde

$$\begin{aligned} v_1 &= (1,1,0) \\ v_2 &= \left(-\frac{5}{7}, \frac{2}{7}, 0\right) \\ v_3 &= (2,-1,1) \end{aligned}$$

y entonces la base β respecto de la cual la forma q tiene la expresión (2.9) es:

$$\beta = \left\{(1,1,0), \left(-\frac{5}{7}, \frac{2}{7}, 0\right), (2,-1,1)\right\}$$

2.2. LA LEY DE LA INERCIA

Considérese el ejemplo presentado en la subsección anterior, en el cual se escribió la forma $q\colon \mathbf{R}^3 \to \mathbf{R}$

$$q(x) = 3x_2^2 + 3x_3^2 + 4x_1x_2 + 4x_1x_3 - 2x_2x_3$$

como una suma de cuadrados de dos maneras distintas. Al observar las expresiones correspondientes

$$q(x) = 4(x_1')^2 + 4(x_2')^2 - 2(x_3')^2 = 7(x_1'')^2 - \frac{4}{7}(x''_2)^2 + 8(x''_3)^2$$

se ve que en ambas aparece la misma cantidad de cuadrados con coeficientes positivos así como con coeficientes negativos: son dos cuadrados, en ambas expresiones, con coeficientes positivos y un cuadrado con coeficiente negativo.

Esto no es una mera coincidencia. Resulta que si una forma cuadrática se presenta como una suma de cuadrados de dos maneras distintas, el número de cuadrados con coeficientes de signo similar en ambas expresiones será siempre el mismo. Esto es lo que se probará a continuación.

TEOREMA 2.2

(LEY DE LA INERCIA PARA FORMAS CUADRÁTICAS.) Sea q: $V \to \mathbf{R}$ una forma cuadrática en el espacio vectorial V de dimensión n. Si β_1 y β_2 son dos bases de V para las que se tienen las expresiones

$$q(x) = \sum_{i=1}^{n} a_i(x_i')^2 = \sum_{i=1}^{n} b_i(x_i'')^2 \qquad (2.11)$$

en donde $(x)_{\beta_1} = (x_1', x_2', \ldots, x_n')$ y $(x)_{\beta_2} = (x_1'', x_2'', \ldots, x_n'')$, entonces el número de coeficientes a_i positivos (negativos) en la primera de estas expresiones es igual al número de coeficientes b_i positivos (negativos) en la segunda.

DEMOSTRACIÓN Se puede suponer (al efectuar una reordenación de los vectores de las bases β_1 y β_2 si fuera necesario) que en la base $\beta_1 = \{v_1, v_2, \ldots, v_n\}$ de V la expresión para la forma cuadrática q es:

$$q(x) = \tilde{a}_1(x_1')^2 + \tilde{a}_2(x_2')^2 + \ldots + \tilde{a}_p(x_p')^2 - \tilde{a}_{p+1}(x_{p+1}')^2 - \ldots - \tilde{a}_{p+q}(x_{p+q}')^2 \qquad (2.12)$$

en donde ahora $\tilde{a}_i$, $i = 1, 2, \ldots, p + q$ son *números positivos* (es decir, se está suponiendo que los cuadrados con coeficientes positivos aparecen todos agrupados en los primeros p lugares, seguidos de los cuadrados con coeficientes negativos). Similarmente, en la base $\beta_2 = \{u_1, u_2, \ldots, u_n\}$ la expresión para q se puede suponer como

$$q(x) = \tilde{b}_1(x_1'')^2 + \tilde{b}_2(x_2'')^2 + \ldots + \tilde{b}_k(x_k'')^2 - \tilde{b}_{k+1}(x_{k+1}')^2 - \ldots - \tilde{b}_{k+m}(x_{k+m}'')^2 \qquad (2.13)$$

en donde también $\tilde{b}_i > 0$, $i = 1, 2, \ldots, k + m$. Se está suponiendo que $p + q \leq n$ y $k + m \leq n$, al aceptar que algunos de los coeficientes a_i y b_i en las expresiones (2.11) sean iguales a cero. Se debe entonces probar que $p = k$ y $q = m$.

Supóngase, para obtener una contradicción, que $p > k \geq 0$.

Sean W_1 y W_2 dos subespacios de V definidos como

$$W_1 = \mathcal{L}(v_1, v_2, \ldots, v_p)$$
$$W_2 = \mathcal{L}(u_{k+1}, u_{k+2}, \ldots, u_n)$$

Obsérvese que

$$\dim W_1 + \dim W_2 = p + n - k > n$$

como

$$\dim (W_1 \cap W_2) = \dim W_1 + \dim W_2 - \dim (W_1 + W_2)$$

y obviamente $\dim(W_1 + W_2) \leq n$, se tiene que $\dim(W_1 \cap W_2) > 0$. Es decir, que $W_1 \cap W_2 \neq \emptyset$.

Tómese entonces un vector $w \in W_1 \cap W_2$. Se puede escribir

$$w = \gamma_1 v_1 + \gamma_2 v_2 + \ldots + \gamma_p v_p \qquad \text{(pues } w \in W_1)$$
$$= \sigma_{k+1}u_{k+1} + \sigma_{k+2}u_{k+2} + \ldots + \sigma_n u_n \qquad \text{(pues } w \in W_2)$$

es decir, que

$$(w)_{\beta_1} = (\gamma_1, \gamma_2, \ldots, \gamma_p, 0, 0, \ldots, 0)$$
$$(w)_{\beta_2} = (0, 0, \ldots, \sigma_{k+1}, \sigma_{k+2}, \ldots, \sigma_n)$$

Según la expresión (2. 12) se tiene que

$$q(w) = \tilde{a}_1(\gamma_1)^2 + \tilde{a}_2(\gamma_2)^2 + \ldots + \tilde{a}_p(\gamma_p)^2$$

de donde se concluye que $q(w) > 0$, pues todos los coeficientes de los cuadrados son positivos y p no es cero.

Por otra parte, según la expresión (2.13) se tiene que

$$q(w) = -\tilde{b}_{k+1}(\sigma_{k+1})^2 - \tilde{b}_{k+2}(\sigma_{k+2})^2 - \ldots - b_{k+m}^2(\sigma_{k+m})^2$$

de donde se concluye que $q(w) \leq 0$, obteniendo así la contradicción deseada.

Se ha mostrado entonces, que $p \leq k$. De un modo similar se puede concluir que $p \geq k$ (al suponer ahora que $p < k$ y llegando a una contradicción análoga a la anterior). Entonces $p = k$ como se quería.

La demostración de que $q = m$ es completamente similar a la anterior y se deja como ejercicio para el lector.

Q.E.D.

Véase un último ejemplo.

EJEMPLO 4 En $\mathbf{R}^3$ considere la forma cuadrática $q: \mathbf{R}^3 \to \mathbf{R}$

$$q(x) = x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 - 2x_2x_3$$

o bien,

$$q(x) = [x_1\, x_2\, x_3]\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Al diagonalizar ortogonalmente la matriz

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

se obtiene la matriz

$$P = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}$$

tal que

$$P^tAP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

o sea que respecto de la base

$$\beta_1 = \left\{\frac{1}{\sqrt{3}}(1, 1, 1), \frac{1}{\sqrt{2}}(1, 0, -1), \frac{1}{\sqrt{6}}(1, -2, 1)\right\}$$

la forma cuadrática q se ve como

$$q(x) = (x_2')^2 + 3(x_3')^2$$

en donde $(x)_{\beta_1} = (x_1', x_2', x_3')$, es decir,

$$x_1' = \frac{x_1 + x_2 + x_3}{\sqrt{3}}$$

$$x_2' = \frac{x_1 - x_3}{\sqrt{2}}$$

$$x_3' = \frac{x_1 - 2x_2 + x_3}{\sqrt{6}}$$

En este caso, se tienen dos coeficientes positivos, un coeficiente cero y ningún coeficiente negativo en la fórmula que expresa a q como una suma de cuadrados.

Por otra parte, al usar la misma idea presentada en la página 616 de ir completando cuadrados en la expresión de $q(x)$ se puede llegar a la fórmula

$$q(x) = (x_1 - x_2)^2 + (x_2 - x_3)^3$$

Es decir, que si se hace

$$x_2'' = x_1 - x_2$$

$$x_3'' = x_2 - x_3$$

se tiene que la forma q se escribe como

$$q(x) = (x_2'')^2 + (x_3'')^2 \tag{2.14}$$

Nuevamente q está escrita como una suma de cuadrados y tal como lo asegura la ley de la inercia, se tienen dos coeficientes positivos, un coeficiente cero y ningún coeficiente negativo.

Se deja como ejercicio al lector verificar que la base β_2 de $\mathbf{R}^3$ respecto de la cual $q(x)$ toma la expresión de suma de cuadrados (2.14) es:

$$\beta_2 = \{(1, 1, 1), (0, -1, -1), (0, 0, -1)\}$$

2.3. FORMAS DEFINIDAS POSITIVAS Y DEFINIDAS NEGATIVAS

DEFINICIÓN 2.3 Se dice que la forma cuadrática $q: V \to \mathbf{R}$ en el espacio vectorial V es una forma

a) definida positiva, si $q(v) > 0 \;\; \forall\, v \in V, v \neq 0$.
b) definida negativa, si $q(v) < 0 \;\; \forall\, v \in V, v \neq 0$.
c) semidefinida positiva, si $q(v) \geq 0 \;\; \forall\, v \in V$.
d) semidefinida negativa, si $q(v) \leq 0 \;\; \forall\, v \in V$.

Por ejemplo, la forma cuadrática $q: \mathbf{R} \to \mathbf{R}$ dada por

$$q(x) = ax^2$$

es definida positiva si $a > 0$, definida negativa si $a < 0$, semidefinida positiva si $a \geq 0$ y semidefinida negativa si $a \leq 0$.

Más generalmente, supóngase que la forma cuadrática $q: V \to \mathbf{R}$ en el espacio vectorial V de dimensión n se presenta (en alguna determinada base β de V) como una suma de cuadrados

$$q(x) = \alpha_1(x_1)^2 + \alpha_2(x_2)^2 + \ldots + \alpha_k(x_k)^2 \quad (k \leq n)$$

en donde $(x)_\beta = (x_1, x_2, \ldots, x_n)$. Es claro entonces que la forma q es definida positiva si, y sólo si $k = n$ y $\alpha_1, \alpha_2, \ldots, \alpha_n$ son números positivos, y definida negativa si, y sólo si $k = n$ y $\alpha_1, \alpha_2, \ldots, \alpha_n$ son números negativos. Además, según la ley de la inercia presentada en la subsección anterior, este hecho no depende de la presentación concreta de la forma q como una suma de cuadrados.

Al usar el corolario del teorema 2.1 se puede afirmar entonces que si A es la matriz asociada de la forma q respecto de alguna base β de V, esta forma será definida positiva (definida negativa) si, y sólo si todos los valores propios de la matriz A son números positivos (negativos). Se deja al lector dar los detalles del argumento que prueba esta afirmación.

EJEMPLO 5

Por ejemplo, la forma q: $\mathbf{R}^2 \to \mathbf{R}$ dada por

$$q(x_1, x_2) = -5x_1^2 - 14x_2^2 + 8x_1x_2$$

es definida negativa, pues la matriz asociada a esta forma respecto de la base canónica de $\mathbf{R}^2$

$$A = \begin{bmatrix} -5 & 4 \\ 4 & -14 \end{bmatrix}$$

tiene por polinomio característico a

$$p(\lambda) = det(A - \lambda I) = det\begin{bmatrix} -5-\lambda & 4 \\ 4 & -14-\lambda \end{bmatrix} = \lambda^2 + 19\lambda + 54$$

cuyas dos raíces son negativas.

EJEMPLO 6

Por otra parte, la forma q: $\mathbf{R}^3 \to \mathbf{R}$ dada por

$$q(x) = x_1^2 + 2x_2^2 + x_3^3 - 2x_1x_2 - 2x_2x_3$$

para la que se vio en la subsección anterior que podría ser escrita como

$$q(x) = (x_2')^2 + (x_3')^2$$

en donde $(x)_\beta = (x_1', x_2', x_3')$ y $\beta = \{(1, 1, 1), (0, -1, -1), (0, 0, -1)\}$, es una forma semidefinida positiva (¿por qué?).

El objetivo principal de esta subsección es presentar un criterio muy útil para decidir cuándo una forma cuadrática es definida positiva (o definida negativa). Para esto, se introducirá primeramente el concepto de "submatriz angular". Sea A la matriz cuadrada de orden n

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

A las submatrices

$$A_1 = [a_{11}], A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \ldots, A_n = A$$

se les llama submatrices angulares de A.

Se denotará por Δ_i al determinante de la submatriz angular A_i de A (entonces $\Delta_n = det\, A$).

TEOREMA 2.3

Sea q: $V \to \mathbf{R}$ una forma cuadrática en el espacio vectorial V de dimensión n. La forma q es definida positiva si, y sólo si la matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12} & a_{22} & \ldots & a_{2n} \\ & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{bmatrix}$$

de la forma q respecto de alguna (cualquiera) base β de V tiene la propiedad de que todos los determinantes de sus submatrices angulares, $\Delta_1, \Delta_2, \ldots, \Delta_n$ son positivos.

DEMOSTRACIÓN La demostración de este resultado se hará por inducción sobre n.

Para $n = 1$ la forma cuadrática q se ve como

$$q(x) = [x]\,[a]\,[x] = ax^2$$

en donde $A = [a]$. El teorema afirma en este caso que la forma q es definida positiva si, y sólo si $a > 0$, lo cual es obviamente cierto.

Obsérvese además que un cambio de base transforma la matriz $A = [a]$ en la matriz P^tAP, en donde P es una matriz cuadrada de orden 1 inversible, esto es, $P = [p]$, $p \neq 0$.

En tal caso,

$$P^tAP = [p]\,[a]\,[p] = p^2a$$

de modo que siendo $a > 0$ y $p \neq 0$, se tiene $p^2 a > 0$, mostrando así que la afirmación del teorema no depende de la base considerada en V, para el caso $n = 1$.

Supóngase entonces válido el teorema para $n = k - 1$ y pruébese para $n = k$.

Se demostrará primero que si la forma q es definida positiva, entonces los determinantes $\Delta_1, \Delta_2, \ldots, \Delta_k$ son positivos.

Se tiene entonces la forma q: $V \to \mathbf{R}$, $dim\, V = k$. Sea β una base (arbitraria pero fija) de V y supóngase que respecto de esta base la matriz asociada a la forma q es

la matriz del enunciado del teorema. Es decir,

$$q(x) = [x]^t_\beta A[x]_\beta$$

se sabe que $q(x)$ puede escribirse como

$$q(x) = \sum_{i,j=1}^{k} a_{ij}x_ix_j = \sum_{i=1}^{k} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$$

en donde $(x)_\beta = (x_1, x_2, \ldots, x_k)$. Se puede reescribir esta última expresión exhibiendo explícitamente a x_k de la siguiente manera:

$$q(x) = \sum_{i,j=1}^{k-1} a_{ij}x_ix_j + 2\sum_{i=1}^{k-1} a_{ik}x_ix_k + a_{kk}x_k^2 \tag{2.15}$$

Sea $\tilde{V}$ un subespacio de V de dimensión $k-1$ y considérese en $\tilde{V}$ la forma cuadrática $\tilde{q}$: $\tilde{V} \to \mathbf{R}$ tal que

$$\tilde{q}(x) = \sum_{i,j=1}^{k-1} a_{ij}x_ix_j$$

es la expresión de la forma $\tilde{q}$, en cierta base $\tilde{\beta}$ de $\tilde{V}$.

Se afirma que la forma $\tilde{q}$ es definida positiva. En efecto, supóngase lo contrario. Existiría entonces un vector $\tilde{x} \in \tilde{V}$, dígase que $(\tilde{x})_{\tilde{\beta}} = (x_1, x_2, \ldots, x_{k-1})$ tal que $\tilde{q}(\tilde{x}) \leq 0$. Considérese el vector $x \in V$ tal que $(x)_\beta = (x_1, x_2, \ldots, x_{k-1}, 0)$. Al evaluar q en x se obtiene, según la expresión (2.15) que $q(x) \leq 0$, lo cual contradice la hipótesis de que la forma q es definida positiva.

Por lo tanto, según la hipótesis de inducción los determinantes de las submatrices angulares de la matriz de la forma $\tilde{q}$ son positivos. Éstos son: $\Delta_1, \Delta_2, \ldots, \Delta_{k-1}$. Resta probar que Δ_k es también positivo.

Como se observa al comienzo de esta subsección, siendo la forma q definida positiva, ésta puede escribirse en alguna base β' de V como una suma de cuadrados con coeficientes positivos. Es decir, $q(x)$ se puede escribir como

$$q(x) = \alpha_1(x_1')^2 + \alpha_2(x_2')^2 + \ldots + \alpha_k(x_k')^2$$

en donde $(x)_{\beta'} = (x_1', x_2', \ldots, x_k')$ y $\alpha_1, \alpha_2, \ldots, \alpha_k$ son números positivos. La matriz asociada a q respecto de la base β' es, pues,

$$D = \begin{bmatrix} \alpha_1 & & 0 \\ & \alpha_2 & \\ 0 & \ddots & \alpha_k \end{bmatrix}$$

la cual está relacionada con la matriz A por medio de la matriz inversible P de cambio de base de β' a β según la fórmula

$$D = P^tAP$$

Al tomar el determinante de las matrices en esta última expresión se obtiene

$$det\ D = det\ (P^tAP) = (det\ P^t)(det\ A)(det\ P) = (det\ A)(det\ P)^2$$

de donde, como $det\ P \neq 0$

$$det\ A = \frac{1}{(det\ P)^2}\, det\ D = \frac{1}{(det\ P)^2}\, \alpha_1\alpha_2 \ldots \alpha_k > 0$$

Se ha probado así que $\Delta_k = det\ A > 0$, como se quería.

Además, este resultado no depende de la base β de V fijada en un comienzo, pues si se cambia a la base β' de V, la matriz asociada a la forma q respecto de esta nueva base β' será $B = P^tAP$, en donde P es la matriz (inversible) de cambio de base de β' a β. Pero en tal caso, $det\ B = (det\ P)^2\ det\ A$. Siendo $det\ A > 0$ y $det\ P \neq 0$, se tiene también que Δ_k para la matriz B es un número positivo pues $\Delta_k = det\ B = (det\ P)^2\ det\ A > 0$.

Se demostrará ahora que si los determinantes $\Delta_1, \Delta_2, \ldots, \Delta_k$ de alguna matriz (arbitraria pero fija) asociada a la forma cuadrática q son positivos, entonces, la forma q es definida positiva. Dígase que A (la matriz del enunciado del teorema) es la matriz que se fija para la forma q.

Considérese nuevamente la forma $\tilde{q}$: $\tilde{V} \to \mathbf{R}$, $dim\ \tilde{V} = k - 1$, tal que respecto de alguna base $\tilde{\beta}$ de $\tilde{V}$ se ve como

$$\tilde{q}(x) = \sum_{i,j=1}^{k-1} a_{ij}x_ix_j$$

Es decir, que la forma $\tilde{q}$ tiene asociada respecto de la base $\tilde{\beta}$ de $\tilde{V}$ a la matriz

$$A_{k-1} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1,k-1} \\ a_{12} & a_{22} & \ldots & a_{2,k-1} \\ & & \cdots & \\ a_{1,k-1} & a_{2,k-1} & \ldots & a_{k-1,k-1} \end{bmatrix}$$

Por la hipótesis de inducción, la forma $\tilde{q}$ es definida positiva pues los determinantes $\Delta_1, \Delta_2, \ldots, \Delta_{k-1}$ son positivos. Entonces, la forma $\tilde{q}$ puede escribirse, respecto de alguna base $\tilde{\beta}'$ de $\tilde{V}$ como

$$\tilde{q}(x) = \alpha_1(x_1')^2 + \alpha_2(x_2')^2 + \ldots + \alpha_{k-1}(x_{k-1}')^2$$

en donde $\alpha_1, \alpha_2, \ldots, \alpha_{k-1}$ son números positivos.

Existe entonces una matriz inversible $\tilde{P}$ de orden $k - 1$ (la matriz de cambio de base de $\tilde{\beta}'$ a $\tilde{\beta}$) tal que

$$\tilde{P}^t A_{k-1} \tilde{P} = \begin{bmatrix} \alpha_1 & & & 0 \\ & \alpha_2 & & \\ & & \ddots & \\ 0 & & & \alpha_{k-1} \end{bmatrix}$$

Sea P la matriz de orden k

$$P = \left[\begin{array}{c|c} \tilde{P} & 0 \\ \hline 0 & 1 \end{array}\right]$$

Obsérvese que P es una matriz inversible, pues $det\ P = det\ \tilde{P} \neq 0$.

Calcúlese el producto P^tAP

$$P^tAP = \left[\begin{array}{c|c} \tilde{P}^t & 0 \\ \hline 0 & 1 \end{array}\right] \left[\begin{array}{c|c} A_{k-1} & \begin{matrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{k-1,k} \end{matrix} \\ \hline a_{1k} a_{2k} \dots a_{k-1,k} & a_{kk} \end{array}\right] \left[\begin{array}{c|c} \tilde{P} & 0 \\ \hline 0 & 1 \end{array}\right]$$

$$= \left[\begin{array}{c|c} \tilde{P}^t A_{k-1} \tilde{P} & \begin{matrix} b_1 \\ b_2 \\ \vdots \\ b_{k-1} \end{matrix} \\ \hline b_1 b_2 \dots b_{k-1} & a_{kk} \end{array}\right] = \left[\begin{array}{cccc|c} \alpha_1 & & & 0 & b_1 \\ & \alpha_2 & & & b_2 \\ & & \ddots & & \vdots \\ 0 & & & \alpha_{k-1} & b_{k-1} \\ \hline b_1 & b_2 & \dots & b_{k-1} & a_{kk} \end{array}\right]$$

en donde $b_1, b_2, \dots, b_{k-1}$, son tales que

$$[a_{1k} \quad a_{2k} \dots a_{k-1,k}]\, \tilde{P} = [b_1 \quad b_2 \dots b_{k-1}]$$

Sea ahora Q la matriz de orden k

$$Q = \left[\begin{array}{c|c} I_{k-1} & \begin{matrix} -b_1/\alpha_1 \\ -b_2/\alpha_2 \\ \vdots \\ -b_{k-1}/\alpha_{k-1} \end{matrix} \\ \hline 0 & 1 \end{array}\right]$$

Obsérvese que Q es una matriz inversible (pues $det\ Q = 1 \neq 0$)

Sea $P' = PQ$. Calcúlese $(P')^tAP'$

$$(P')^tAP' = (PQ)^tA(PQ) = Q^t(P^tAP)Q$$

$$= \left[\begin{array}{ccc|c} & I_{k-1} & & 0 \\ \hline -\dfrac{b_1}{\alpha_1} & \dots & -\dfrac{b_{k-1}}{\alpha_{k-1}} & 1 \end{array}\right] \left[\begin{array}{cccc|c} \alpha_1 & & & & b_1 \\ & \alpha_2 & & & b_2 \\ & & \ddots & & \vdots \\ & & & \alpha_{k-1} & b_{k-1} \\ \hline b_1 & b_2 & \dots & b_{k-1} & a_{kk} \end{array}\right] \left[\begin{array}{c|c} I_{k-1} & \begin{matrix} -b_1/\alpha_1 \\ \vdots \\ -b_{k-1}/\alpha_{k-1} \end{matrix} \\ \hline 0 & 1 \end{array}\right]$$

$$= \begin{bmatrix} \alpha_1 & & & & 0 \\ & \alpha_2 & & & \\ & & \ddots & & \\ & & & \alpha_{k-1} & \\ 0 & & & & b \end{bmatrix}$$

en donde

$$b = a_{kk} - \frac{b_1^2}{\alpha_1} - \frac{b_2^2}{\alpha_2} - \dots - \frac{b_{k-1}^2}{\alpha_{k-1}}$$

Como P' es un producto de dos matrices inversibles P' es también inversible. Sea β' la base de V tal que P' es la matriz de cambio de base de β' a β. Se tiene entonces que la forma cuadrática q se escribe respecto de la base β' como

$$q(x) = [x]^t_{\beta'} (P')^tAP'[x]_{\beta'} = \alpha_1(x_1')^2 + \alpha_2(x_2')^2 + \dots + \alpha_{k-1}(x_{k-1}')^2 + b(x_k')^2$$

en donde $(x)_{\beta'} = (x_1', x_2', \dots, x_k')$. Como $\alpha_1, \alpha_2, \dots, \alpha_{k-1}$ son números positivos, de la expresión anterior se sigue que la forma q es definida positiva si, y sólo si b es un número positivo.

Pero

$$\alpha_1 \alpha_2 \dots \alpha_{k-1}\, b = det(P')^tAP' = (det\ P')^2 A$$

de donde

$$b = \frac{1}{\alpha_1 \alpha_2 \dots \alpha_{k-1}} (det\ P')^2\, det\ A$$

Por hipótesis $\Delta_k = det\ A > 0$. Entonces $b > 0$ y con esto se concluye que la forma q es definida positiva.

Esto concluye la demostración del teorema.

Q.E.D.

EJEMPLO 7 Por ejemplo, considérese la forma q: $\mathbf{R}^3 \to \mathbf{R}$ dada por

$$q(x) = 5x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3$$

La matriz de q respecto de la base canónica de $\mathbf{R}^3$ es:

$$A = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 1 & -1 \\ -1 & -1 & 3 \end{bmatrix}$$

Los determinantes Δ_1, Δ_2 y Δ_3 para esta matriz son:

$$\Delta_1 = det[5] = 5 > 0 \qquad \Delta_2 = det\begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} = 1 > 0$$

$$\Delta_3 = det\begin{bmatrix} 5 & 2 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 3 \end{bmatrix} = 1 > 0$$

Como los tres determinantes Δ_1, Δ_2 y Δ_3 son positivos, se concluye, por el teorema 2.3, que la forma q es una forma definida positiva.

Un corolario del teorema 2.3 que proporciona un criterio para identificar formas definidas negativas es el siguiente:

COROLARIO

Con la notación del teorema 2.3, la forma q: $V \to \mathbf{R}$ es una forma definida negativa si, y sólo si $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \ldots$ Es decir, los determinantes $\Delta_1, \Delta_2, \ldots, \Delta_n$ alternan sus signos, comenzando con $\Delta_1 < 0$.

DEMOSTRACIÓN Es claro que la forma q: $V \to \mathbf{R}$

$$q(x) = [x]_\beta^t A[x]_\beta$$

es definida negativa si, y sólo si la forma $-q$: $V \to \mathbf{R}$

$$(-q)(x) = [x]_\beta^t (-A)[x]_\beta$$

es definida positiva. Según el teorema 2.3 se debe tener entonces, que los determinantes

$$\Delta_1 = det[-a_{11}],\ \Delta_2 = det\begin{bmatrix} -a_{11} & -a_{12} \\ -a_{12} & -a_{22} \end{bmatrix}, \ldots, \Delta_n = det(-A)$$

deben ser positivos. Es decir, que

$$\Delta_1 = det[-a_{11}] = -det[a_{11}] < 0$$

$$\Delta_2 = det\begin{bmatrix} -a_{11} & -a_{12} \\ -a_{12} & -a_{22} \end{bmatrix} = det\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} > 0$$

etcétera, lo que prueba el corolario.

Q.E.D.

EJEMPLO 8 Por ejemplo, la forma q: $\mathbf{R}^3 \to \mathbf{R}$

$$q(x) = -19x_1^2 - 6x_2^2 - 3x_3^2 - 20x_1x_2 - 2x_1x_3 - 4x_2x_3$$

tiene asociada la siguiente matriz respecto de la base canónica de $\mathbf{R}^3$:

$$A = \begin{bmatrix} -19 & -10 & -1 \\ -10 & -6 & -2 \\ -1 & -2 & -3 \end{bmatrix}$$

Se tiene entonces que

$$\Delta_1 = det[-19] = -19 < 0$$

$$\Delta_2 = det\begin{bmatrix} -19 & -10 \\ -10 & -6 \end{bmatrix} = 14 > 0$$

$$\Delta_3 = det\begin{bmatrix} -19 & -10 & -1 \\ -10 & -6 & -2 \\ -1 & -2 & -3 \end{bmatrix} = -38 < 0$$

Se concluye entonces, por el corolario anterior, que esta forma es definida negativa.

EJERCICIOS (SECCIÓN 2, CAPÍTULO 7)

1. Sea V un espacio vectorial de dimensión n. Sea S el conjunto de formas bilineales simétricas en V. Demuestre que S es un subespacio de $L_2(V, \mathbf{R})$, el espacio vectorial de formas bilineales en V. Determine la dimensión de S.
2. Para cada una de las formas cuadráticas q: $\mathbf{R}^2 \to \mathbf{R}$ dadas a continuación, determine la forma bilineal simétrica f: $\mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ que genera a q
   a) $q(x) = x_1^2 + x_2^2$
   b) $q(x) = 2x_1^2$
   c) $q(x) = 6x_2^2$
   d) $q(x) = -5x_1x_2$
   e) $q(x) = 3x_1^2 + 10x_1x_2 + x_2^2$

3. Escriba cada una de las formas cuadráticas $q: \mathbf{R}^3 \to \mathbf{R}$ dadas a continuación como $q(x) = [x]^t_\beta A[x]_\beta$, en donde β es la base canónica de $\mathbf{R}^3$ y A es una matriz simétrica de orden 3
   a) $q(x) = x_1^2 + x_2^2 + x_3^2$
   b) $q(x) = x_1^2 + 6x_1x_2 + x_2^2 + 8x_1x_3 + x_3^2$
   c) $q(x) = -x_1^2 + 10x_2^2 + 7x_3^2 + 5x_1x_2 + 6x_1x_3 - 10x_2x_3$
4. Repita el ejercicio anterior si β es la base de $\mathbf{R}^3$ dada por

$$\beta = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$$

5. Sea V un espacio vectorial con el producto interno $(\cdot \mid \cdot)$. Compruebe que la forma bilineal $f: V \times V \to \mathbf{R}$

$$f(x, y) = (x \mid y)$$

es simétrica. ¿Cuál es la forma cuadrática q que genera f? Si V es de dimensión finita y β es una base de V, escriba la forma cuadrática q generada por f como $q(x) = [x]^t_\beta A[x]_\beta$. (A la matriz A se le llama "matriz del producto interno" $(\cdot \mid \cdot)$ en relación de la base β). Describa explícitamente la matriz A en los siguientes casos:
   a) $V = \mathbf{R}^3$, $(x \mid y) = x_1y_1 + 6x_2y_2 + 10x_3y_3$, β es la base canónica de $\mathbf{R}^3$.
   b) $V = P_2$, $(p \mid q) = p(0)\,q(0) + p(1)\,q(1) + p(2)\,q(2)$ (véase ejercicio 8, sección 1, capítulo 5), β es la base canónica de P_2.
6. Sea V un espacio vectorial de dimensión n y sea β una base de V. Se define el *rango* de una forma cuadrática $q: V \to \mathbf{R}$ como el rango de la matriz de la forma bilineal simétrica que genera a q respecto a la base β. Demuestre que esta definición *no depende* de la base β fijada para V.
7. Determine el rango de cada una de las formas cuadráticas del ejercicio 2.
8. Determine el rango de cada una de las formas cuadráticas del ejercicio 3.
9. Demuestre que la forma cuadrática $q: \mathbf{R}^2 \to \mathbf{R}$ dada por

$$q(x) = ax_1^2 + 2bx_1x_2 + cx_2^2$$

tiene rango 2 si, y sólo si $ac - b^2 \neq 0$.
10. Sea V un espacio vectorial de dimensión n. Sea $q: V \to \mathbf{R}$ una forma cuadrática en V. Pruebe que la forma q tiene rango n si, y sólo si $\lambda = 0$ *no* es un valor propio de la matriz asociada a la forma bilineal simétrica $f: V \times V \to R$ que genera a q respecto de alguna (cualquiera) base β de V.
11. Considere la forma cuadrática $q: \mathbf{R}^n \to \mathbf{R}$ dada por

$$q(x) = (a_1x_1 + a_2x_2 + \ldots + a_nx_n)^2, \quad a_1 \neq 0$$

Escriba esta forma como $q(x) = [x]^t_\beta A[x]_\beta$ en donde A es una matriz simétrica de orden n y
   a) β es la base canónica de $\mathbf{R}^n$
   b) β es la base $\{v_1, v_2, \ldots, v_n\}$ en donde

$$v_1 = (1/a_1, 0, 0, \ldots, 0)$$

$$v_2 = (-a_2/a_1, 1, 0, \ldots, 0)$$
$$v_3 = (-a_3/a_1, 0, 1, \ldots, 0)$$
$$\vdots$$
$$v_n = (-a_n/a_1, 0, 0, \ldots, 1)$$

Con la perspectiva del "cambio de variables", ¿a qué equivale escribir la forma q respecto de la base β del inciso b)? Determine el rango de la forma q.

② 12. Considere la forma cuadrática $q: \mathbf{R}^n \to \mathbf{R}$, $n \neq 1$, dada por

$$q(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - m)^2 \text{ en donde } m = \frac{1}{n}\sum_{j=1}^{n} x_j$$

Escriba esta forma como $q(x) = [x]^t_\beta A[x]_\beta$ en donde β es la base canónica de $\mathbf{R}^n$ y A es una matriz simétrica de orden n. Determine el rango de q.

Escriba cada una de las formas cuadráticas de los ejercicios 13-26 como una suma de cuadrados de *dos maneras distintas*. En cada caso: a) describa explícitamente la base β respecto de la cual q se escribe como $q(x) = [x]^t_\beta D[x]_\beta$, D una matriz diagonal; b) describa explícitamente las "ecuaciones de cambio de variables", esto es, las ecuaciones que relacionan las coordenadas $x_1, x_2, \ldots, x_n$ de un vector x con las nuevas coordenadas $x_1', x_2', \ldots, x_n'$ del vector x en la nueva base β [del inciso (a)]; c) verifique que se satisface la ley de la inercia.

13. $q(x_1, x_2) = 2x_1^2 - x_2^2 + x_1x_2$
14. $q(x_1, x_2) = x_1^2 - 4x_2^2 + 8x_1x_2$
15. $q(x_1, x_2, x_3) = 7x_1^2 + 7x_2^2 + 7x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3$
16. $q(x_1, x_2, x_3) = 3x_1^2 + 4x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_2x_3$
17. $q(x_1, x_2, x_3) = 3x_1^2 + 6x_2^2 + 3x_3^2 - 4x_1x_2 - 8x_1x_3 - 4x_2x_3$
18. $q(x_1, x_2, x_3) = 11x_1^2 + 5x_2^2 + 2x_3^2 + 16x_1x_2 + 4x_1x_3 - 20x_2x_3$
19. $q(x_1, x_2, x_3) = 6x_1^2 + 3x_2^2 + 3x_3^2 + 4x_1x_2 + 4x_1x_3 - 8x_2x_3$
20. $q(x_1, x_2, x_3) = 7x_1^2 + 5x_2^2 + 3x_3^2 - 8x_1x_2 + 8x_2x_3$
21. $q(x_1, x_2, x_3, x_4) = 5x_1^2 + 5x_2^2 + 5x_3^2 + 5x_4^2 - 10x_1x_2 + 2x_1x_3 + 6x_1x_4 + 6x_2x_3 + 2x_2x_4 - 10x_3x_4$
22. $q(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + 2x_1x_2 - 2x_1x_4 - 2x_2x_3 + 2x_3x_4$
23. $q(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2 - 2x_1x_2 + 6x_1x_3 - 4x_1x_4 - 4x_2x_3 + 6x_2x_4 - 2x_3x_4$
24. $q(x_1, x_2, x_3, x_4) = 8x_1x_3 + 2x_1x_4 + 2x_2x_3 + 8x_2x_4$
25. $q(x_1, x_2, x_3, x_4, x_5) = 4x_1^2 + x_2^2 + 5x_3^2 - 4x_4^2 + x_5^2 - 4x_1x_2 + 12x_4x_5$
26. $q(x_1, x_2, \ldots, x_6) = 3x_1^2 - 3x_2^2 + 4x_3^2 - 4x_4^2 + 4x_5^2 + x_6^2 + 8x_1x_2 - 6x_3x_4 + 4x_5x_6$

④ 27. Escribir la forma cuadrática $q: \mathbf{R}^n \to \mathbf{R}$

$$q(x) = \sum_{i=1}^{x} x_i^2 + \sum_{i<j}^{n} x_ix_j$$

como una suma de cuadrados de dos maneras distintas. Verifique que se cumple la ley de la inercia.

28. Sea V un espacio vectorial con el producto interno $(\cdot|\cdot)$. Considere la forma cuadrática $q: V \to \mathbf{R}$ dada por

$$q(v) = (v \mid v)$$

(Véase ejercicio 5.) Demuestre que q es una forma definida positiva.

② 29. Considere la forma cuadrática $q: \mathbf{R}^2 \to \mathbf{R}$ dada por

$$q(x) = ax_1^2 + 2bx_1x_2 + cx_2^2$$

Demuestre que si la forma q es definida positiva, entonces los coeficientes a y c son positivos. Por medio de un ejemplo demuestre que la afirmación recíproca es falsa.

30. Sea V un espacio vectorial de dimensión n. Sea $q: V \to \mathbf{R}$ una forma cuadrática en V. Compruebe que si esta forma es definida positiva entonces el rango de q es n. ¿Es cierta esta afirmación si se cambia "definida positiva" por "definida negativa"?, ¿por "semidefinida positiva"?, ¿por "semidefinida negativa"?

31. Demuestre por medio de un ejemplo, que la recíproca de la afirmación del ejercicio anterior es falsa.

32. Demuestre que las siguientes formas cuadráticas son definidas positivas:

a) $q: \mathbf{R}^2 \to \mathbf{R}$, $q(x) = 2x_1^2 + 8x_2^2 - 3x_1x_2$

b) $q: \mathbf{R}^3 \to \mathbf{R}$, $q(x) = 17x_1^2 + 14x_2^2 + 14x_3^2 - 4x_1x_2 - 4x_1x_3 - 8x_2x_3$

c) $q: \mathbf{R}^3 \to \mathbf{R}$, $q(x) = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_{1_3} - 8x_2x_3$

33. Demuestre que las siguientes formas cuadráticas son definidas negativas:

a) $q: \mathbf{R}^2 \to \mathbf{R}$, $q(x) = -2x_1^2 - x_2^2 - x_1x_2$

b) $q: \mathbf{R}^3 \to \mathbf{R}$, $q(x) = -6x_1^2 - 5x_2^2 - 7x_3^2 + 4x_1x_2 - 4x_1x_3$

c) $q: \mathbf{R}^3 \to \mathbf{R}$, $q(x) = -5x_1^2 - 6x_2^2 - 4x_3^2 + 4x_1x_2 + 4x_1x_3$

34. Determine los valores de α para los cuales la forma $q: \mathbf{R}^3 \to \mathbf{R}$

$$q(x) = 5x_1^2 + x_2^2 + \alpha x_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3$$

es definida positiva.

35. Compruebe que para ningún valor de α la forma $q: \mathbf{R}^3 \to \mathbf{R}$

$$q(x) = x_1^2 + 4x_2^2 + x_3^2 + 2\alpha x_1x_2 + 10x_1x_3 + 6x_2x_3$$

puede ser definida positiva.

36. Considere la matriz simétrica

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

a) Demuestre que los valores propios de A se encuentran en el intervalo $[-1, 5]$.

b) Considere ahora la forma cuadrática $q: \mathbf{R}^3 \to \mathbf{R}$ dada por

$$q(x) = [x]_\beta^t (A - \lambda I) [x]_\beta$$

en donde β es la base canónica de $\mathbf{R}^3$. Pruebe que esta forma es definida negativa si $\lambda > 5$ y definida positiva si $\lambda < -1$.

② 37. El ejercicio anterior es un caso particular del siguiente hecho general: si los valores propios de la matriz simétrica A de orden n se encuentran en el intervalo $[a, b]$, entonces la forma cuadrática $q: \mathbf{R}^n \to \mathbf{R}$ dada por

$$q(x) = [x]_\beta^t (A - \lambda I) [x]_\beta$$

(β es la base canónica de $\mathbf{R}^n$, dígase) es definida negativa si $\lambda > b$ y definida positiva si $\lambda < a$. Demuéstrelo. ¿Es cierta la afirmación recíproca?

38. Se dice que una matriz simétrica A de orden n es definida positiva (definida negativa, semidefinida positiva, semidefinida negativa) si la forma cuadrática $q: \mathbf{R}^n \to \mathbf{R}$

$$q(x) = [x]_\beta^t A [x]_\beta$$

lo es. Demuestre que una condición necesaria y suficiente para que una matriz simétrica A sea definida positiva (definida negativa) es que todos sus valores propios sean números positivos (negativos, respectivamente).

39. Compruebe que una matriz simétrica A es definida positiva (definida negativa) si, y sólo si los determinantes de las submatrices angulares de A $\Delta_1, \Delta_2, \ldots, \Delta_n$ son todos positivos (tienen signos alternados, comenzando con $\Delta_1 < 0$, respectivamente).

40. Demuestre que si A y B son matrices del mismo orden definidas positivas (negativas), entonces $A + B$ es también una matriz definida positiva (negativa, respectivamente).

41. ¿Cierto o falso? El conjunto de matrices definidas positivas (negativas) de orden n es un subespacio de $M_{n \times n}$.

④ 42. Compruebe que si A es una matriz definida positiva, existe una matriz M (del mismo orden que A) inversible tal que $A = M^tM$. ¿Es esta matriz M única? Verifique este resultado con la matriz

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{bmatrix}$$

[*Sugerencia*: Existe una matriz ortogonal P tal que $D = P^tAP$ es una matriz diagonal con los valores propios de A (que son números positivos) en su diagonal principal. Escriba D como $D = E^2$. Continúe. . .]

43. (Una aplicación a la determinación de extremos locales de funciones de varias variables).

Se considerarán en este ejercicio funciones *suaves* $f: \mathbf{R}^n \to \mathbf{R}$. En cálculo de funciones de varias variables se demuestra que una condición *necesaria* para que la función $f: \mathbf{R}^n \to \mathbf{R}$ tenga un extremo local en el punto $x_0 \in \mathbf{R}^n$, es que en ese punto se anulen todas las derivadas parciales de f, esto es,

$$\frac{\partial f}{\partial x_i}(x_0) = 0, \quad i = 1, 2, \ldots, n$$

En tal caso, se dice que x_0 es un *punto crítico* de f. Considere ahora la matriz de derivadas parciales de segundo orden $A = (a_{ij})_{i,j=1,\ldots,n}$ en donde

$$a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0), \quad i, j = 1, 2, \ldots, n$$

Ésta es una matriz simétrica de orden n. Se puede demostrar la validez del siguiente

criterio, que establece condiciones *suficientes* para la existencia de extremos locales de la función f: si x_0 es un punto crítico de f: $\mathbf{R}^n \to \mathbf{R}$, entonces,

a) Si la matriz $A = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)\right)_{i,j=1,\ldots,n}$ es definida positiva, f tiene un *mínimo local* en x_0.

b) Si la matriz A es definida negativa, f tiene un *máximo local* en x_0.

c) Si la matriz A tiene rango n, pero no es definida positiva ni definida negativa, f tiene un *punto de ensilladura* en x_0.

d) Si la matriz A tiene un rango menor que n, no se puede afirmar nada acerca de la naturaleza del punto x_0.

Por ejemplo, considere la función f: $\mathbf{R}^2 \to \mathbf{R}$ dada por

$$f(x_1, x_2) = x_1^3 + x_2^3 - 6x_1x_2$$

Los puntos críticos se obtienen al resolver el sistema

$$\frac{\partial f}{\partial x_1} = 3x_1^2 - 6x_2 = 0$$

$$\frac{\partial f}{\partial x_2} = 3x_2^2 - 6x_1 = 0$$

Éstos son entonces $p_1 = (0, 0)$ y $p_2 = (2, 2)$. Las derivadas parciales de segundo orden son:

$$\frac{\partial^2 f}{\partial x_1^2} = 6x_1, \ \frac{\partial^2 f}{\partial x_1 \partial x_2} = -6 = \frac{\partial^2 f}{\partial x_2 \partial x_1}, \ \frac{\partial^2 f}{\partial x_2^2} = 6x_2$$

En el punto p_1 la matriz A de parciales de segundo orden es:

$$A = \begin{bmatrix} 0 & -6 \\ -6 & 0 \end{bmatrix}$$

Los valores propios de esta matriz son $\lambda_1 = 6$ y $\lambda_2 = -6$. Se está entonces en el caso c). Se concluye que la función f tiene un punto de ensilladura en p_1, que vale $f(0, 0) = 0$.

En el punto p_2 la matriz B de parciales de segundo orden es:

$$B = \begin{bmatrix} 12 & -6 \\ -6 & 12 \end{bmatrix}$$

cuyos valores propios son $\lambda_1 = 6$ y $\lambda_2 = 18$ (ambos positivos), o bien cuyos determinantes angulares son $\Delta_1 = 12$ $\Delta_2 = 108$ (también ambos positivos). Se está entonces en el caso a). Se concluye que la función f tiene un mínimo local en el punto p_2 que vale $f(2, 2) = -8$.

1) Demuestre que la función $f(x_1, x_2) = x_1^2 + x_2^2$ tiene un mínimo local en el punto (0, 0). La superficie que representa esta función es un paraboloide elíptico (véase pág. 679).

2) Compruebe que la función $f(x_1, x_2) = x_1^2 - x_2^2$ tiene un punto de ensilladura en el punto (0, 0). La superficie que representa esta función es un paraboloide hiperbólico (véase pág. 679).

3) Determine la naturaleza de los puntos críticos de la función

$$f(x_1, x_2) = x_1x_2\sqrt{1 - x_1^2/3 - x_2^2/3}$$

4) Determine la naturaleza de los puntos críticos de la función

$$f(x_1, x_2) = \frac{1}{2}x_1x_2 + (47 - x_1 - x_2)\left(\frac{x_1}{3} + \frac{x_2}{4}\right)$$

5) Determine la naturaleza de los puntos críticos de la función

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1 - 2x_3$$

3. PARÁBOLAS, ELIPSES, HIPÉRBOLAS, ETC.

En esta sección será de interés estudiar la *ecuación general de segundo grado*

$$f(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0 \tag{3.1}$$

Uno de los objetivos será el de clasificar el conjunto de puntos en el plano

$$\{(x, y) \in \mathbf{R}^2 \,|\, f(x, y) = 0\} \ *$$

También interesará describir gráficamente el lugar geométrico que representa la ecuación (3.1).

Esta sección tiene un sabor predominantemente geométrico, a diferencia de las dos secciones anteriores, en donde seguía existiendo, como en todo el libro, el sabor algebraico. En realidad el estudio que en esta sección se emprende no se encuentra propiamente identificado dentro del álgebra lineal. Éste queda más bien encuadrado dentro del ámbito de la geometría (analítica). Resulta, sin embargo, que las herramientas proporcionadas por el álgebra lineal permiten tener un acercamiento bastante elegante y eficiente a la solución del problema de clasificación de los lugares geométricos descritos por la ecuación (3.1) que, como se había dicho, dar solución a este problema es uno de los objetivos en esta sección.

El primer paso que se dará en esta dirección será demostrar que siempre es posible efectuar una transformación de coordenadas de modo que la ecuación (3.1) puede verse, en el sistema coordenado transformado, como una ecuación en la que no aparezca el término del producto de las coordenadas xy.

Más concretamente, se verá que existe una transformación T: $\mathbf{R}^2 \to \mathbf{R}^2$ (de coordenadas)

$$T(x, y) = (x', y')$$

tal que en el nuevo sistema de coordenadas $x'y'$ la ecuación (3.1) puede verse como

$$A'x'^2 + C'y'^2 + 2D'x' + 2E'y' + F' = 0$$

---

*En toda esta sección se considerará a $\mathbf{R}^2$ como un espacio vectorial con el producto interno canónico.

Considérese para esto la forma cuadrática

$$q(x, y) = Ax^2 + 2Bxy + Cy^2 = [x \quad y]\begin{bmatrix} A & B \\ B & C \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

La matriz M de la forma q respecto de la base canónica $\beta = \{e_1, e_2\}$ de $\mathbf{R}^2$ es entonces,

$$M = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

Se sabe, por el corolario del teorema 2.1 de la sección anterior, que existe una base β' de $\mathbf{R}^2$ respecto de la cual la expresión de la forma q se ve como

$$q(x, y) = \lambda_1 x'^2 + \lambda_2 y'^2$$

en donde $(x, y)_{\beta'} = (x', y')$ y los coeficientes λ_1 y λ_2 son precisamente los valores propios de la matriz M.

Aún más, del corolario mencionado anteriormente se puede deducir que la base β' es una base *ortonormal* (al igual que la base β —la canónica de $\mathbf{R}^2$—), formada por vectores propios asociados a cada uno de los valores propios λ_1 y λ_2.

Sea P la matriz de cambio de base de β' a β. Obsérvese que P es una matriz ortogonal (¿por qué?).

Considérese la transformación $T: \mathbf{R}^2 \to \mathbf{R}^2$ dada por

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = P^t\begin{bmatrix} x \\ y \end{bmatrix}$$

Siendo P la matriz de cambio de base de β' a β, se tiene que $P^t = P^{-1}$ (pues P es ortogonal) es la matriz de cambio de la base de β a β', de modo que la imagen del vector (x, y) $(=(x, y)_\beta)$ bajo la transformación T [esto es, $T(x, y)$] es la matriz de coordenadas de este vector en la base β'. Llámese (x', y') a las coordenadas del vector (x, y) es la nueva base β'. Se tiene, pues, que

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = P^t\begin{bmatrix} x \\ y \end{bmatrix} = P^{-1}\begin{bmatrix} x \\ y \end{bmatrix}$$

Entonces,

$$\begin{bmatrix} x \\ y \end{bmatrix} = P\begin{bmatrix} x' \\ y' \end{bmatrix}$$

Se puede contemplar esta última expresión como las *ecuaciones de transformación* (del sistema coordenado original xy al nuevo sistema $x'y'$). Éstas son, pues,

$$x = ax' + by'$$

$$y = cx' + dy'$$

en donde $u_1 = (a, c)$, $u_2 = (b, d)$ son los vectores de la nueva base ortonormal β'.

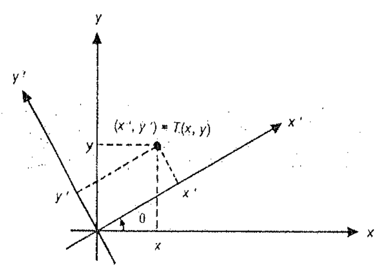
Al sustituir estas expresiones en la ecuación (3.1) se llega a una ecuación del tipo

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2D'x' + 2E'y' + F' = 0 \tag{3.2}$$

en donde D', E' y F' son unos nuevos coeficientes.

Si se considera un nuevo sistema coordenado $x'y'$, en el cual los ejes x' y y' se encuentran en las direcciones de los vectores u_1 y u_2 de la base β', se puede ver resumido el análisis anterior diciendo que *al referir la curva que representa la ecuación* (3.1) *al nuevo sistema coordenado* $x'y'$,* *esta ecuación se transforma en la ecuación* (3.2).

Obsérvese, por otra parte, que siendo P una matriz ortogonal, la matriz P^t también lo es y entonces la transformación T descrita anteriormente es una *transformación ortogonal* (teorema 5.1 capítulo 5). Según el análisis que sigue al teorema 5.3 del capítulo 5, se puede concluir entonces que el nuevo sistema $x'y'$ se obtiene del sistema original xy *girando éste en cierto ángulo* θ.



Se dice entonces que la ecuación (3.2) se ha obtenido de la ecuación (3.1) por medio de una *rotación de los ejes coordenados*.

EJEMPLO 1

Considere por ejemplo la ecuación

$$3x^2 - 2xy + 3y^2 + 2x - 4y + 1 = 0 \tag{3.3}$$

Se tiene en este caso que la matriz M es:

$$M = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

cuyo polinomio característico es:

$$p(\lambda) = det(M - \lambda I) = det\begin{bmatrix} 3-\lambda & -1 \\ -1 & 3-\lambda \end{bmatrix} = \lambda^2 - 6\lambda + 8 = (\lambda - 4)(\lambda - 2)$$

y entonces los valores propios de M son $\lambda_1 = 4$ y $\lambda_2 = 2$.

*A los ejes $x'y'$ se les llama *ejes principales* de la curva representada por (3.1).

Es fácil ver que los vectores

$$u_1 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \qquad u_2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

son vectores propios (ortonormales) correspondientes a cada uno de estos valores propios (respectivamente). Si se considera la base $\beta' = \{u_1, u_2\}$ [respecto de la cual la forma cuadrática $q\colon \mathbb{R}^2 \to \mathbb{R}$, $q(x, y) = 3x^2 - 2xy + 3y^2$ se debe ver como $q(x, y) = 4x'^2 + 2y'^2$ en donde $(x, y)_{\beta'} = (x', y')$], se tiene que la matriz de cambio de base de β' a β es:

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

de modo que las ecuaciones de transformación por rotación de ejes son:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

o sea,

$$x = \frac{1}{\sqrt{2}}(x' + y')$$

$$y = \frac{1}{\sqrt{2}}(-x' + y')$$

Al sustituir estas expresiones en la ecuación (3.3) y al simplificar se obtiene

$$4x'^2 + 2y'^2 + \frac{6}{\sqrt{2}}x' - \frac{2}{\sqrt{2}}y' + 1 = 0$$

que es una ecuación del tipo (3.2).

El próximo objetivo será hacer una simplificación más en la ecuación (3.2) de modo que se llegue a establecer así la *ecuación canónica* de la curva que representa la ecuación (3.1) —y entonces poder identificar el lugar geométrico que representa esta ecuación—. En este punto, se tendrán que considerar dos casos:

CASO 1. $\lambda_1\lambda_2 \neq 0$

En este caso se puede escribir la ecuación (3.2) como

$$\lambda_1\left(x'^2 + \frac{2D'}{\lambda_1}x'\right) + \lambda_2\left(y'^2 + \frac{2E'}{\lambda_2}y'\right) + F' = 0$$

o bien, al completar cuadrados

$$\lambda_1\left(x' + \frac{D'}{\lambda_1}x'\right)^2 + \lambda_2\left(y' + \frac{E'}{\lambda_2}\right)^2 + G = 0$$

en donde $G = F' - D'^2/\lambda_1 - E'^2/\lambda_2$. Haga las sustituciones

$$x'' = x' + \frac{D'}{\lambda_1}$$

$$y'' = y' + \frac{E'}{\lambda_2} \tag{3.4}$$

Se tiene entonces que la ecuación (3.2) se transforma en

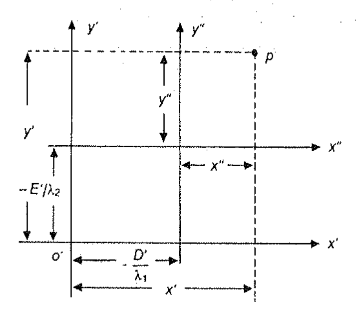
$$\lambda_1 x''^2 + \lambda_2 y''^2 + G = 0 \tag{3.5}$$

Sobre esta ecuación ya es fácil decidir sobre el lugar geométrico que representa.

A continuación, se verán las diferentes posibilidades de lugares geométricos que puede representar esta ecuación.

Observe que el efecto geométrico de las ecuaciones (3.4) corresponde a una *traslación del origen de coordenadas*: el origen del sistema $x'y'$ se traslada al punto $\left(-\frac{D'}{\lambda_1}, -\frac{E'}{\lambda_2}\right)$ obteniendo así el sistema $x''y''$. Se dice entonces que la ecuación (canónica) (3.5) se ha obtenido de la ecuación (3.2) por medio de una traslación de ejes coordenados.

Esquemáticamente se tiene



SUBCASO **1.1.** $\lambda_1\lambda_2 > 0$ y $G\lambda_1 < 0$

Se puede escribir la ecuación (3.5) como

$$\frac{x''^2}{-G/\lambda_1} + \frac{y''^2}{-G/\lambda_2} = 1$$

Como $G\lambda_1 < 0$ (y por lo tanto, $G\lambda_2 < 0$), se puede escribir esta última expresión como

$$\frac{x''^2}{a^2} + \frac{y''^2}{b^2} = 1$$

para ciertas constantes no nulas a y b. Ésta es la ecuación canónica de una *elipse*.

SUBCASO **1.2.** $\lambda_1\lambda_2 > 0$ y $G = 0$
En este caso la ecuación (3.5) se ve como

$$\lambda_1 x''^2 + \lambda_2 y''^2 = 0$$

Esta ecuación se satisface si, y sólo si $x'' = y'' = 0$. Entonces en este caso la ecuación (3.5) [y por lo tanto la (3.1)] representa *un punto* [en el sistema x' y' este punto es $\left(-\frac{D'}{\lambda_1}, -\frac{E'}{\lambda_2}\right)$ el origen de coordenadas del sistema $x''y''$].

SUBCASO **1.3.** $\lambda_1\lambda_2 > 0$ y $G\lambda_1 > 0$.
Al igual que en el subcaso 1.1, se escribe la ecuación (3.5) como

$$\frac{x''^2}{-G/\lambda_1} + \frac{y''^2}{-G/\lambda_2} = 1$$

Por ser $G\lambda_1 > 0$ (y también $G\lambda_2 > 0$), se ve que el primer miembro de esta expresión es un número no positivo. Por lo tanto, no existen valores algunos de x'' y y'' que la satisfagan. En este caso entonces la ecuación (3.1) *no representa lugar geométrico alguno* en el plano (se dice que su gráfica es un conjunto vacío).

SUBCASO **1.4.** $\lambda_1\lambda_2 < 0$ y $G \neq 0$.
Al escribir la ecuación (3.5) como

$$\frac{x''^2}{-G/\lambda_1} + \frac{y''^2}{-G/\lambda_2} = 1$$

se ve que siendo en este caso λ_1 y λ_2 de signos opuestos, esta última ecuación puede escribirse como

$$\frac{x''^2}{a^2} - \frac{y''^2}{b^2} = 1$$

o bien,

$$\frac{y''^2}{a^2} - \frac{x''^2}{b^2} = 1$$

para ciertas constantes no nulas a y b.
Éstas son las ecuaciones canónicas de una *hipérbola*.

SUBCASO **1.5.** $\lambda_1\lambda_2 < 0$ y $G = 0$
Suponga por ejemplo que $\lambda_1 > 0$. Se puede entonces escribir la ecuación (3.5) como

$$a^2x''^2 - b^2y''^2 = 0$$

para ciertas constantes no nulas a y b. El primer miembro de esta última expresión

es una diferencia de cuadrados que puede reescribirse como

$$(ax'' - by'')(ax'' + by'') = 0$$

y entonces la ecuación se satisface si, y sólo si

$$ax'' - by'' = 0$$
$$ax'' + by'' = 0$$

Éstas son las ecuaciones de *dos rectas que se cortan* en el origen (del sistema $x''y''$).

CASO 2. $\lambda_1\lambda_2 = 0$
Sin pérdida de generalidad, se puede suponer que $\lambda_1 = 0$. Se tienen entonces dos subcasos.

SUBCASO **2.1.** $\lambda_1 = 0$, $\lambda_2 \neq 0$, $D' \neq 0$.
La ecuación (3.2) tiene en este caso la forma

$$\lambda_2 y'^2 + 2D'x' + 2E'y' + F' = 0$$

que se puede reescribir como

$$\lambda_2\left(y'^2 + \frac{2E'}{\lambda_2}y' + \frac{E'^2}{\lambda_2^2}\right) + 2D'\left(x' + \frac{F'}{2D'} - \frac{E'^2}{2\lambda_2 D'}\right) = 0$$

o bien,

$$\lambda_2\left(y' + \frac{E'}{\lambda_2}\right)^2 + 2D'\left(x' + \frac{F'}{2D'} - \frac{E'^2}{2\lambda_2 D'}\right) = 0$$

Al realizar

$$y'' = y' + \frac{E'}{\lambda_2}$$

$$x'' = x' + \frac{F'}{2D'} - \frac{E'^2}{2\lambda_2 D'}$$

la ecuación (3.2) se ve como

$$\lambda_2 y''^2 + 2D'x'' = 0$$

que es una ecuación del tipo

$$y''^2 - 4px'' = 0, \quad p = -\frac{D'}{2\lambda_2}$$

Es decir, en este caso la ecuación (3.1) representa una *parábola*.

SUBCASO 2.2. $\lambda_1 = 0, \lambda_2 \neq 0, D' = 0.$
En este caso, la ecuación (3.2) se ve como

$$\lambda_2 y'^2 + 2E'y' + F' = 0$$

o bien,

$$\lambda_2\left(y' + \frac{E'}{\lambda_2}\right)^2 + F' - \frac{E'^2}{\lambda_2} = 0$$

Al hacer

$$y'' = y' + \frac{E'}{\lambda_2}$$

$$x'' = x'$$

$$c = F' - \frac{E'^2}{\lambda_2}$$

queda finalmente que si se realiza una traslación del origen de coordenadas del sistema $x'y'$ al punto $(0, -E'/\lambda_2)$, la ecuación (3.2) se transforma en

$$\lambda_2 y''^2 + c = 0 \tag{3.6}$$

de donde

a) si $\lambda_2 c < 0$, la ecuación (3.6) se ve como

$$y'' = \pm\sqrt{-\frac{c}{\lambda_2}}$$

que son las ecuaciones de *dos rectas paralelas distintas.*

b) si $c = 0$, la ecuación (3.6) se ve como

$$\lambda_2 y''^2 = 0$$

que se satisface si, y sólo si $y'' = 0$. Ésta es la ecuación del eje x''. Es decir, en este caso la ecuación (3.7) —y por tanto, la (3.1)— representa *una recta* (o dos rectas paralelas confundidas en el eje x'').

c) Si $\lambda_2 c > 0$, la ecuación (3.7) se ve como

$$y''^2 = -\frac{c}{\lambda_2}$$

Siendo el segundo miembro de esta expresión un número negativo, es claro que en este caso la gráfica de la ecuación (3.7) es un *conjunto vacío.*

NOTA: No se considera el caso en el que $\lambda_1 = \lambda_2 = 0$, pues éste provendría de la ecuación

$$2Dx + 2Ey + F = 0$$

la cual no es una ecuación de *segundo* grado y obviamente su lugar geométrico es una recta en el plano. En efecto, si $\lambda_1 = \lambda_2 = 0$, la matriz

$$M = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

sería semejante a la matriz cero (¿por qué?). Entonces, la matriz M misma sería cero, es decir $A = B = C = 0$, mostrando así que los únicos términos no nulos en la ecuación (3.1) son los términos lineales.

Se tienen entonces clasificados todos los tipos de lugares geométricos que puede representar la ecuación general de segundo grado (3.1). Obsérvese que en esta primera clasificación, para poder determinar el tipo de lugar geométrico representado por (3.1) es necesario primeramente hacer la simplificación de esta ecuación por rotación de ejes, estableciendo así la ecuación (3.2); en base a ella fue que se hizo el análisis de las diferentes posibilidades de valores de sus coeficientes. El siguiente objetivo será hacer más práctica y eficiente la identificación de los lugares geométricos correspondientes *trabajando directamente con los coeficientes de la ecuación* (3.1).

Pero antes véanse algunos ejemplos.

EJEMPLO 2

Considere la ecuación

$$4x^2 - 24xy + 11y^2 + 56x - 58y + 95 = 0 \tag{3.7}$$

La matriz M asociada a la forma cuadrática $q: \mathbf{R}^2 \to \mathbf{R}$, $q(x, y) = 4x^2 - 24xy + 11y^2$ respecto de la base canónica es:

$$M = \begin{bmatrix} 4 & -12 \\ -12 & 11 \end{bmatrix}$$

cuyo polinomio característico es:

$$p(\lambda) = det(M - \lambda I) = det\begin{bmatrix} 4-\lambda & -12 \\ -12 & 11-\lambda \end{bmatrix} = (\lambda - 20)(\lambda + 5)$$

y entonces, $\lambda_1 = 20$ y $\lambda_2 = -5$ son los valores propios de M.

Una base ortonormal de $\mathbf{R}^2$ formada por vectores propios de M es $\beta' = \{u_1, u_2\}$ en donde

$$u_1 = \left(-\frac{3}{5}, \frac{4}{5}\right) \qquad u_2 = \left(\frac{4}{5}, \frac{3}{5}\right)$$

de modo que la matriz P de cambio de base de β' a la **base canónica** de $\mathbf{R}^2$ es:

$$P = \begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix}$$

Las ecuaciones de transformación por rotación de ejes son entonces,

$$\begin{bmatrix} x \\ y \end{bmatrix} = P\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix}$$

es decir,

$$x = -\frac{3}{5}x' + \frac{4}{5}y'$$

$$y = \frac{4}{5}x' + \frac{3}{5}y'$$

Al sustituir estas expresiones en (3.7) y al simplificar se obtiene

$$20x'^2 - 5y'^2 + 80x' + 10y' + 95 = 0 \qquad (3.8)$$

que es una ecuación del tipo (3.2) con $\lambda_1 = 20$, $\lambda_2 = -5$, $D' = 40$, $E' = 5$, $F' = 95$.

Como $\lambda_1\lambda_2 = -100 \neq 0$, uno se encuentra en el caso 1. Se debe entonces trasladar el origen de coordenadas al punto

$$\left(-\frac{D'}{\lambda_1}, -\frac{E'}{\lambda_2}\right) = (-2, 1)$$

de modo que al sustituir, según (3.4)

$$x' = x'' - 2$$

$$y' = y'' + 1$$

en la ecuación (3.8) y al simplificar se obtiene

$$20x''^2 - 5y''^2 + 20 = 0$$

Se tiene entonces $\lambda_1\lambda_2 = -100 < 0$ y $G = 20 \neq 0$. Éste es el subcaso 1.4. Se concluye entonces que la ecuación (3.7) representa una hipérbola cuya ecuación canónica en el sistema $x''y''$ (el sistema original xy girado y trasladado) es:

$$\frac{y''^2}{4} - x''^2 = 1$$

Véase otro ejemplo.

EJEMPLO 3 Considere la ecuación

$$x^2 + 4xy + 4y^2 - 2x - 4y - 3 = 0 \qquad (3.9)$$

La matriz M es:

$$M = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

cuyo polinomio característico es:

$$p(\lambda) = det(M - \lambda I) = det\begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix} = \lambda(\lambda - 5)$$

y entonces los valores propios de M son $\lambda_1 = 0$ y $\lambda_2 = 5$. Los vectores $u_1 = \frac{1}{\sqrt{5}}(-2, 1)$ y $u_2 = \frac{1}{\sqrt{5}}(1, 2)$ constituyen una base ortonormal de $\mathbf{R}^2$ formada por vectores propios correspondientes a λ_1 y λ_2, respectivamente. La matriz P es entonces

$$P = \begin{bmatrix} -2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

Al hacer

$$\begin{bmatrix} x \\ y \end{bmatrix} = P\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} -2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix}$$

o sea,

$$x = -\frac{2}{\sqrt{5}}x' + \frac{1}{\sqrt{5}}y'$$

$$y = \frac{1}{\sqrt{5}}x' + \frac{2}{\sqrt{5}}y'$$

y al sustituir en la ecuación (3.9) queda después de simplificar

$$5y'^2 - \frac{10}{\sqrt{5}}y' - 3 = 0$$

que es una ecuación del tipo (3.2) —simplificada por rotación de ejes coordenados— con $\lambda_1 = 0$, $\lambda_2 = 5$, $D' = 0$, $E' = -5/\sqrt{5}$, $F' = -3$. Se está entonces en el subcaso 2.2, pues $\lambda_1 = 0$, $\lambda_2 = 5 \neq 0$ y $D' = 0$.

Al escribir la última expresión como

$$5y''^2 - 4 = 0$$

en donde

$$y'' = y' - 1/\sqrt{5}$$

se ve que se trata de dos rectas paralelas, a saber $y'' = \pm 2/\sqrt{5}$. En este caso, entonces, la ecuación (3.9) representa dos rectas paralelas.

3.1. REDUCCIÓN DE LA ECUACIÓN GENERAL DE SEGUNDO GRADO A LA FORMA CANÓNICA

Como se había dicho ya anteriormente uno se ocupará ahora de hacer más eficiente la simplificación de la ecuación general de segundo grado por rotación y traslación de los ejes coordenados, y por consiguiente, la identificación del lugar geométrico que tal ecuación representa.

Se parte entonces de la ecuación general de segundo grado

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0 \qquad (3.1)$$

El primer paso que se ha realizado con el fin de simplificar esta ecuación consiste en referirla a un nuevo sistema coordenado que se obtiene del sistema original girando éste un cierto ángulo θ. Con esto se logra una ecuación del tipo

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2D'x' + 2E'y' + F' = 0 \qquad (3.2)$$

Esta ecuación se obtiene pues de (3.1) simplificando ésta "por rotación de ejes".

Como ya se vio, lo que se hace para obtener (3.2) a partir de (3.1) es realizar la sustitución

$$\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x' \\ y' \end{bmatrix}$$

en la ecuación (3.1), en donde P es la matriz de cambio de base de β a la base canónica de $\mathbf{R}^2$, siendo β una base ortonormal de $\mathbf{R}^2$ formada por vectores propios de la matriz M

$$M = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

Al disponer ya de la ecuación (3.2), se realizaba en ella la sustitución

$$x' = x'' + h$$

$$y' = y'' + k$$

con h y k escogidas "convenientemente" con el fin de lograr una nueva simplifición en la ecuación. Esto equivalía a trasladar el origen de coordenadas (del sistema $x'y'$) al punto (h, k).

Se tiene, pues, que la ecuación (3.1) sufre dos modificaciones (simplificaciones) al realizar en ella rotaciones y traslaciones de ejes. Esquemáticamente se tiene

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

↓ ROTACIÓN

$$A'x'^2 + 2B'x'y' + C'y'^2 + 2D'x' + 2E'y' + F' = 0$$

↓ TRASLACIÓN

$$A''x''^2 + 2B''x''y'' + C''y''^2 + 2D''x'' + 2E''y'' + F'' = 0$$

Se entiende que la última de estas ecuaciones posee una estructura más simple que la inicial.

Un hecho muy interesante sobre el cual se quiere llamar la atención es que existen ciertas cantidades relacionadas con los coeficientes de estas ecuaciones, que conservan su valor a través de las simplificaciones realizadas en la ecuación original. Se dice que estas cantidades son *invariantes por rotación y traslación de ejes*.

EJEMPLO 4 Considere el ejemplo que ya se había estudiado en la página 643. Se tiene

$$4x^2 - 24xy + 11y^2 + 56x - 58y + 95 = 0$$

↓ ROTACIÓN

$$20x'^2 + 0 \cdot x'y' - 5y'^2 + 80x' + 10y' + 95 = 0$$

↓ TRASLACIÓN

$$20x''^2 + 0 \cdot x''y'' - 5y''^2 + 0 \cdot x'' + 0 \cdot y'' + 20 = 0$$

Observe que

a) $A + C = 4 + 11 = 15$

$A' + C' = A'' + C'' = 20 - 5 = 15$

b) $det\begin{bmatrix} A & B \\ B & C \end{bmatrix} = det\begin{bmatrix} 4 & -12 \\ -12 & 11 \end{bmatrix} = -100$

$det\begin{bmatrix} A' & B' \\ B' & C' \end{bmatrix} = det\begin{bmatrix} A'' & B'' \\ B'' & C'' \end{bmatrix} = det\begin{bmatrix} 20 & 0 \\ 0 & -5 \end{bmatrix} = -100$

c) $det\begin{bmatrix} A & B & D \\ B & C & E \\ C & E & F \end{bmatrix} = det\begin{bmatrix} 4 & -12 & 28 \\ -12 & 11 & -29 \\ 28 & -29 & 95 \end{bmatrix} = -2000$

$det\begin{bmatrix} A' & B' & D' \\ B' & C' & E' \\ D' & E' & F' \end{bmatrix} = det\begin{bmatrix} 20 & 0 & 40 \\ 0 & -5 & 5 \\ 40 & 5 & 95 \end{bmatrix} = -2000$

$$\det\begin{bmatrix} A'' & B'' & D'' \\ B'' & C'' & E'' \\ D'' & E'' & F'' \end{bmatrix} = \det\begin{bmatrix} 20 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 20 \end{bmatrix} = -2000$$

Éste es un hecho que se explotará posteriormente para simplificar eficientemente una ecuación general de segundo grado.

Ahora se enuncia y se demuestra en general.

TEOREMA 3.1

Sea la ecuación general de segundo grado

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

Las cantidades

$$\omega = A + C$$

$$\mu = \det\begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

$$\nu = \det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix}$$

se mantienen invariantes si en la ecuación se realiza una rotación y/o una traslación de ejes coordenados.

DEMOSTRACIÓN Supóngase primeramente que se realiza una traslación del origen de coordenadas al punto (h, k). Es decir, que en la ecuación (3.1) se hace la sustitución

$$x = x' + h$$

$$y = y' + k$$

Se tiene entonces,

$$A(x' + h)^2 + 2B(x' + h)(y' + k) + C(y' + k)^2 + 2D(x' + h) + 2E(y' + k) + F = 0$$

o bien,

$$Ax'^2 + 2Bx'y' + Cy'^2 + 2(Ah + Bk + D)x' + 2(Bh + Ck + E)y' + Ah^2 + 2Bhk + Ck^2 + 2Dh + 2Ek + F = 0$$

que es una ecuación (transformada por traslación) del tipo

$$A'x'^2 + 2B'x'y' + C'y'^2 + 2D'x' + 2E'y' + F' = 0$$

en donde

$$A' = A$$

$$B' = B$$

$$C' = C$$

$$D' = Ah + Bk + D$$

$$E' = Bh + Ck + E$$

$$F' = Ah^2 + 2Bhk + Ck^2 + 2Dh + 2Ek + F$$

Calcúlense las cantidades ω, μ y ν para esta ecuación transformada por traslación.

Se tiene

$$\omega = A' + C' = A + C$$

$$\mu = \det\begin{bmatrix} A' & B' \\ B' & C' \end{bmatrix} = \det\begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

y así entonces los valores de ω y μ se mantuvieron invariantes. Se verá que ν se mantiene también sin variación.

$$\nu = \det\begin{bmatrix} A' & B' & D' \\ B' & C' & E' \\ D' & E' & F' \end{bmatrix}$$

$$= \det\begin{bmatrix} A & B & Ah + Bk + D \\ B & C & Bh + Ck + E \\ Ah + Bk + D & Bh + Ck + E & Ah^2 + 2Bhk + Ck^2 + 2Dh + 2Ek + F \end{bmatrix}$$

En la matriz del determinante de esta última expresión se realizan las siguientes transformaciones: sustitúyase la tercera línea por ella misma menos h veces la primera y menos k veces la segunda. El valor del determinante no se altera y queda entonces

$$\nu = \det\begin{bmatrix} A & B & Ah + Bk + D \\ B & C & Bh + Ck + D \\ D & E & Dh + Ek + F \end{bmatrix}$$

Si se sustituye ahora la tercera columna por ella misma menos h veces la primera y menos k veces la segunda se obtiene

$$\nu = \det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix}$$

mostrando así que el valor de ν no se altera por traslación de ejes en la ecuación (3.1).

Véase ahora que los valores de ω y μ no se alteran si se efectúa una rotación de ejes en la ecuación general de segundo grado original.

Obsérvese que los valores de ω y μ dependen solamente de los coeficientes de la forma cuadrática $q: \mathbf{R}^2 \to \mathbf{R}$

$$q(x, y) = Ax^2 + 2Bxy + Cy^2 = [x \quad y]\begin{bmatrix} A & B \\ B & C \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Efectuar una rotación en los ejes coordenados equivale a realizar la sustitución [en la ecuación (3.1)].

$$\begin{bmatrix} x \\ y \end{bmatrix} = P\begin{bmatrix} x' \\ y' \end{bmatrix}$$

en donde P es una matriz ortogonal (véase el análisis que sigue al teorema 5.3 del capítulo 5). En tal caso la forma cuadrática q se transforma en

$$q(x, y) = [x' \; y'] \; P^t \begin{bmatrix} A & B \\ B & C \end{bmatrix} P \begin{bmatrix} x' \\ y' \end{bmatrix} = [x' \; y'] \begin{bmatrix} A' & B' \\ B' & C' \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix}$$

El polinomio característico de la matriz

$$\begin{bmatrix} A' & B' \\ B' & C' \end{bmatrix}$$

es:

$$p(\lambda) = det\begin{bmatrix} A' - \lambda & B' \\ B' & C' - \lambda \end{bmatrix} = \lambda^2 - (A' + C')\lambda + (A'C' - B'^2)$$

Al ser esta matriz semejante a la matriz

$$\begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

(pues al ser P ortogonal se tiene

$$\begin{bmatrix} A' & B' \\ B' & C' \end{bmatrix} = P^{-1}\begin{bmatrix} A & B \\ B & C \end{bmatrix} P)$$

ambas poseen el mismo polinomio característico. Entonces,

$$p(\lambda) = det\begin{bmatrix} A - \lambda & B \\ B & C - \lambda \end{bmatrix} = \lambda^2 - (A + C)\lambda + (AC - B^2)$$
$$= \lambda^2 - (A' + C')\lambda + (A'C' - B'^2)$$

de donde

$$\omega = A + C = A' + C'$$
$$\mu = det\begin{bmatrix} A & B \\ B & C \end{bmatrix} = AC - B^2 = A'C' - B'^2 = det\begin{bmatrix} A' & B' \\ B' & C' \end{bmatrix}$$

mostrando así la invariancia de ω y μ por rotación en los ejes coordenados.

Por último, para ver que ν se mantiene también invariante por rotación, considérese la forma cuadrática $q: \mathbf{R}^3 \to \mathbf{R}$

$$q(x, y, z) = Ax^2 + 2Bxy + Cy^2 + 2Dxz + 2Eyz + Fz^2 \qquad (3.10)$$
$$= [x \; y \; z]\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Como ya se había dicho, efectuar una rotación en los ejes coordenados en la ecuación (3.1) equivale a hacer la sustitución

$$\begin{bmatrix} x \\ y \end{bmatrix} = P\begin{bmatrix} x' \\ y' \end{bmatrix} \qquad (3.11)$$

en donde P es una matriz ortogonal.

Sea $\tilde{P}$ la matriz de orden 3

$$\tilde{P} = \left[\begin{array}{c|c} P & 0 \\ \hline 0 & 1 \end{array}\right]$$

Obsérvese que $\tilde{P}$ es también una matriz ortogonal, pues

$$\tilde{P}^t\tilde{P} = \left[\begin{array}{c|c} P^t & 0 \\ \hline 0 & 1 \end{array}\right]\left[\begin{array}{c|c} P & 0 \\ \hline 0 & 1 \end{array}\right] = \left[\begin{array}{c|c} I_2 & 0 \\ \hline 0 & 1 \end{array}\right] = I_3$$

En la expresión(3.10) para la forma cuadrática q hágase la sustitución

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \tilde{P}\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

Se obtiene

$$q(x, y, z) = [x' \; y' \; z'] \; \tilde{P}^t \begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} \tilde{P} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

Obsérvese que

$$\tilde{P}^t \begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} \tilde{P} = \left[\begin{array}{c|c} P^t \begin{bmatrix} A & B \\ B & C \end{bmatrix} P & P^t \begin{bmatrix} D \\ E \end{bmatrix} \\ \hline [D \quad E] \, P & F \end{array}\right] \tag{3.12}$$

Por otra parte, cuando se efectúa la sustitución (3.11) en la ecuación (3.1), se obtiene la ecuación transformada (por rotación)

$$A'x'^2 + 2B'x'y' + C'y'^2 + 2D'x' + 2E'y' + F' = 0 \tag{3.13}$$

en donde los nuevos coeficientes $A', B', \ldots, F'$ están relacionados con los coeficientes originales $A, B, \ldots, F$ según

$$\begin{bmatrix} A' & B' \\ B' & C' \end{bmatrix} = P^t \begin{bmatrix} A & B \\ B & C \end{bmatrix} P$$

$$\begin{bmatrix} D' \\ E' \end{bmatrix} = P^t \begin{bmatrix} D \\ E \end{bmatrix}$$

$$F' = F$$

Esto se puede ver si se escribe la ecuación (3.1) como

$$[x \; y] \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + 2[D \; E] \begin{bmatrix} x \\ y \end{bmatrix} + F = 0$$

y se realiza en ésta la sustitución (3.11) para obtener así la ecuación (3.13).

Entonces, según (3.12) se tiene que

$$\tilde{P}^t \begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} \tilde{P} = \begin{bmatrix} A' & B' & D' \\ B' & C' & E' \\ D' & E' & F' \end{bmatrix}$$

de donde, por ser $\tilde{P}$ una matriz ortogonal, se concluye que las matrices

$$\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} A' & B' & D' \\ B' & C' & E' \\ D' & E' & F' \end{bmatrix}$$

son semejantes. En particular, ambas tienen el mismo determinante. Esto es precisamente lo que se quería probar.

Q.E.D.

Se procederá ahora a hacer el análisis del lugar geométrico que representa una ecuación general de segundo grado, usando el teorema que se ha demostrado.

Se tiene, pues, que la ecuación (3.1)

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

es transformada, por medio de una rotación en los ejes coordenados en

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2D'x' + 2E'y' + F' = 0$$

en donde λ_1 y λ_2 son los valores propios de la matriz

$$M = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

Debido a la invariancia de la cantidad μ del teorema anterior, se debe tener entonces que

$$\mu = det \begin{bmatrix} A & B \\ B & C \end{bmatrix} = det \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \lambda_1 \lambda_2$$

Por lo tanto, los dos casos que se habían considerado en la clasificación de los lugares geométricos descritos por la ecuación (3.1) —los casos $\lambda_1\lambda_2 \neq 0$ y $\lambda_1\lambda_2 = 0$— corresponden, respectivamente, a los casos en que $\mu \neq 0$ y $\mu = 0$.

Considérese el primer caso, $\mu \neq 0$.

Se vio ya que en estas circunstancias ($\mu \neq 0$), por medio de una traslación de los ejes de coordenadas se puede establecer la ecuación simplificada

$$\lambda_1 x''^2 + \lambda_2 y''^2 + G = 0$$

Según lo mostrado en el teorema (3.1), el determinante ν se debe mantener invariante, es decir que

$$\nu = det \begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = det \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & G \end{bmatrix} = \lambda_1 \lambda_2 G$$

de donde

$$G = \frac{\nu}{\lambda_1\lambda_2} = \frac{\nu}{\mu}$$

Se puede decir entonces que la forma canónica de la ecuación (3.1) en este caso ($\mu \neq 0$) es:

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \frac{\nu}{\mu} = 0 \tag{3.14}$$

Analícense los subcasos correspondientes al caso 1.

Subcaso **1.1.** $\lambda_1\lambda_2 > 0$ y $G\lambda_1 < 0$ (la gráfica es una elipse)

En términos de las cantidades ω, μ y ν, este subcaso corresponde a que $\mu = \lambda_1\lambda_2 > 0$ y $\nu = G\mu \neq 0$. Además como

$$\begin{aligned} \omega\nu &= (\lambda_1 + \lambda_2)G\lambda_1\lambda_2 \\ &= \lambda_1^2\lambda_2 G + \lambda_2^2\lambda_1 G \end{aligned}$$

y λ_1 y λ_2 tienen el mismo signo, se ve que $G\lambda_1 < 0$ si, y sólo si $\omega\nu < 0$. En resumen, este subcaso corresponde a que $\mu > 0$, $\nu \neq 0$ y $\omega\nu < 0$.

Subcaso **1.2.** $\lambda_1\lambda_2 > 0$ y $G = 0$ (la gráfica es un punto).

En este subcaso se tiene $\mu > 0$ y $\nu = G\mu = 0$.

Subcaso **1.3.** $\lambda_1\lambda_2 > 0$ y $G\lambda_1 > 0$ (la gráfica es un conjunto vacío).

Al igual que en el subcaso 1.1 se puede ver fácilmente que este subcaso corresponde a $\mu > 0$, $\nu \neq 0$ y $\omega\nu > 0$.

Subcaso **1.4.** $\lambda_1\lambda_2 < 0$ y $G \neq 0$ (la gráfica es una hipérbola).

Este subcaso corresponde a que $\mu < 0$ y $\nu = G\mu \neq 0$.

Subcaso **1.5.** $\lambda_1\lambda_2 < 0$ y $G = 0$ (la gráfica son dos rectas que se cortan).

Este subcaso corresponde a $\mu < 0$ y $\nu = G\mu = 0$.

Pase ahora al caso 2, $\lambda_1\lambda_2 = 0$, es decir, el caso en el que $\mu = 0$.

Subcaso **2.1.** $\lambda_1 = 0$, $\lambda_2 \neq 0$, $D' \neq 0$ (la gráfica es una parábola).

En este subcaso se llega a establecer la ecuación canónica de la ecuación (3.1) como

$$\lambda_2 y''^2 + 2D'x'' = 0$$

Debido a la invariancia de la cantidad ν se debe tener entonces que

$$\nu = det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = det\begin{bmatrix} 0 & 0 & D' \\ 0 & \lambda_2 & 0 \\ D' & 0 & 0 \end{bmatrix} = -D'^2\lambda_2$$

Como $\lambda_2 \neq 0$ y $D' \neq 0$, se tiene que $\nu \neq 0$. Este subcaso corresponde pues a que $\nu \neq 0$.

Subcaso **2.2.** $\lambda_1 = 0$, $\lambda_2 \neq 0$, $D' = 0$ (la gráfica es: dos rectas paralelas, una recta o un conjunto vacío).

En este subcaso se estableció la ecuación canónica de la ecuación (3.1) como

$$\lambda_2 y''^2 + c = 0$$

Por la invariancia de ν se debe tener que

$$\nu = det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = det\begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & c \end{bmatrix} = 0$$

Este subcaso corresponde, pues, a $\nu = 0$.

En el siguiente cuadro se resume el análisis anterior. Con él se cumplió el objetivo que se plantea al inicio de esta sección de clasificar los lugares geométricos que representa una ecuación general de segundo grado

---

ECUACIÓN: $Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$

$$\omega = A + C, \quad \mu = det\begin{bmatrix} A & B \\ B & C \end{bmatrix}, \quad \nu = det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix}$$

LUGAR GEOMÉTRICO QUE REPRESENTA

$\mu \neq 0$	$\mu > 0$	$\nu \neq 0$	$\mu\nu < 0$.	**Una elipse**
			$\mu\nu > 0$.	**Ningún lugar geométrico**
		$\nu = 0$		**Un punto**
	$\mu < 0$	$\nu \neq 0$		**Una hipérbola**
		$\nu = 0$		**Dos rectas que se cortan**
$\mu = 0$	$\nu \neq 0$			**Una parábola**
	$\nu = 0$			**Dos rectas paralelas, o una recta, o ningún lugar geométrico**

Introdúzcase una nueva terminología respecto de esta clasificación. Cuando en la ecuación general de segundo grado se tiene $\mu \neq 0$, se dice que la curva (el lugar geométrico) que representa tal ecuación es una *curva con centro*. Caso contrario ($\mu = 0$) se dice que es una *curva sin centro*. Más aún, cuando $\mu > 0$ se dice que la *curva es de tipo elíptico* y cuando $\mu < 0$ se dice que la *curva es de tipo hiperbólico*. A los casos en que la ecuación general de segundo grado no representa una

parábola, una elipse o una hipérbola se les llama *casos degenerados*. Se tiene, pues, la siguiente versión del cuadro anterior:

ECUACIÓN: $Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$

LUGAR GEOMÉTRICO QUE REPRESENTA

- **Curva con centro** ($\mu \neq 0$)
  - **De tipo elíptico** ($\mu > 0$)
    - **Una elipse** ($\omega\nu < 0$)
    - **Casos degenerados** ($\omega\nu \geq 0$)
  - **De tipo hiperbólico** ($\mu < 0$)
    - **Una hipérbola** ($\nu \neq 0$)
    - **Caso degenerado** ($\nu = 0$)
- **Curva sin centro** ($\mu = 0$)
  - **Una parábola** ($\nu \neq 0$)
  - **Casos degenerados** ($\nu = 0$)

Véanse algunos ejemplos.

EJEMPLO 5 Para la ecuación

$$3x^2 - 4xy - 4y^2 + 16x + 16y - 12 = 0 \tag{3.15}$$

se tiene

$$\mu = det\begin{bmatrix} A & B \\ B & C \end{bmatrix} = det\begin{bmatrix} 3 & -2 \\ -2 & -4 \end{bmatrix} = -16 < 0$$

$$\nu = det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = det\begin{bmatrix} 3 & -2 & 8 \\ -2 & -4 & 8 \\ 8 & 8 & -12 \end{bmatrix} = 0$$

Por lo tanto, la ecuación (3.15) representa dos rectas que se cortan.

EJEMPLO 6 Para la ecuación

$$52x^2 - 72xy + 73y^2 - 104x + 72y - 48 = 0 \tag{3.16}$$

se tiene

$$\mu = det\begin{bmatrix} A & B \\ B & C \end{bmatrix} = det\begin{bmatrix} 52 & -36 \\ -36 & 73 \end{bmatrix} = 2500 > 0$$

$$\nu = det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = det\begin{bmatrix} 52 & -36 & -52 \\ -36 & 73 & 36 \\ -52 & 36 & -48 \end{bmatrix} = -250000 \neq 0$$

$$\omega = A + C = 52 + 73 = 125$$

$$\omega\nu = (125)(-250000) < 0$$

Se tiene entonces que la ecuación (3.16) representa una elipse. Para establecer la ecuación canónica de esta elipse sólo se necesita, según la ecuación (3.14), los valores propios de la matriz M

$$M = \begin{bmatrix} 52 & -36 \\ -36 & 73 \end{bmatrix}$$

(y los valores de μ y ν que ya se calcularon). El polinomio característico de M es:

$$p(\lambda) = det\begin{bmatrix} 52-\lambda & -36 \\ -36 & 73-\lambda \end{bmatrix} = \lambda^2 - 125\lambda + 2500 = (\lambda - 25)(\lambda - 100)$$

en donde $\lambda_1 = 25$ y $\lambda_2 = 100$. Entonces la ecuación canónica de la elipse que representa (3.16) es:

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \frac{\nu}{\mu} = 25x''^2 + 100y''^2 - \frac{250000}{2500} = 0$$

o sea,

$$\frac{x''^2}{4} + y''^2 = 1$$

EJEMPLO 7 Para la ecuación

$$x^2 - 10xy + y^2 - 10x + 2y + 13 = 0 \tag{3.17}$$

se tiene

$$\mu = det\begin{bmatrix} A & B \\ B & C \end{bmatrix} = det\begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix} = -24 < 0$$

$$\nu = det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = det\begin{bmatrix} 1 & -5 & -5 \\ -5 & 1 & 1 \\ -5 & 1 & 13 \end{bmatrix} = -288 \neq 0$$

La ecuación (3.17) representa entonces una hipérbola. Halle su ecuación canónica.

Se tiene

$$p(\lambda) = det\begin{bmatrix} 1-\lambda & -5 \\ -5 & 1-\lambda \end{bmatrix} = \lambda^2 - 2\lambda - 24 = (\lambda - 6)(\lambda + 4)$$

de donde $\lambda_1 = 6$ y $\lambda_2 = -4$. Por lo tanto, según la ecuación (3.14) la ecuación canónica de esta hipérbola es:

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \frac{\nu}{\mu} = 6x''^2 - 4y''^2 + \frac{-228}{-24} = 0$$

o sea,

$$\frac{y''^2}{3} - \frac{x''^2}{2} = 1$$

EJEMPLO 8 Por último, para la ecuación

$$9x^2 + 24xy + 16y^2 + 132x - 74y - 81 = 0 \quad (3.18)$$

se tiene

$$\mu = det\begin{bmatrix} A & B \\ B & C \end{bmatrix} = det\begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix} = 0$$

$$\nu = det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = det\begin{bmatrix} 9 & 12 & 66 \\ 12 & 16 & -37 \\ 66 & -37 & -81 \end{bmatrix} = -140625 \neq 0$$

Por lo tanto, la ecuación representa una parábola. Vea cuál es su ecuación canónica.

Se había visto que en este caso (véase subcaso 2.1, pág. 641) la ecuación canónica de la parábola es:

$$\lambda_2 y''^2 + 2D'x'' = 0$$

Halle λ_2. Se tiene

$$p(\lambda) = det\begin{bmatrix} 9-\lambda & 12 \\ 12 & 16-\lambda \end{bmatrix} = \lambda(\lambda - 25)$$

Entonces ($\lambda_1 = 0$ y) $\lambda_2 = 25$.

Por otra parte,

$$\nu = -D'^2\lambda_2$$

de donde

$$D' = \sqrt{-\frac{\nu}{\lambda_2}} = \sqrt{5625} = 75*$$

La ecuación canónica de la parábola es entonces

$$25y''^2 + 150x'' = 0$$

o sea,

$$y''^2 + 6x'' = 0$$

3.2. CONSTRUCCIÓN DE LA GRÁFICA DE UNA ECUACIÓN GENERAL DE SEGUNDO GRADO

Hasta este momento se han dedicado todos los esfuerzos a *identificar* el tipo de lugar geométrico que representa una ecuación general de segundo grado, así como a establecer la ecuación canónica de tal lugar geométrico. Sin embargo, es claro que resulta deseable disponer de una gráfica de él. Sobre esto se trabajará en esta subsección.

Como ya se ha visto anteriormente, la ecuación canónica de la curva que representa una ecuación general de segundo grado se obtiene realizando en esta última simplificaciones primeramente por rotación de ejes coordenados (eliminando así el término con el producto de las coordenadas) y posteriormente por traslación de los ejes (eliminando así el —o los— términos de primer grado).

Considérese la primera de estas transformaciones, es decir, la rotación de los ejes coordenados. Se verá ahora cómo determinar las *direcciones* de los nuevos ejes del sistema girado.

Según el análisis que se presentó al inicio de esta sección, la transformación por rotación de ejes consiste en escribir la ecuación original (3.1) en una nueva base (ortonormal) de $\mathbb{R}^2$: la base constituida por vectores propios —ortonormales— de la matriz

$$M = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

Sea $\beta' = \{u_1, u_2\}$ tal base. Véase cómo determinar las direcciones de los vectores u_1 y u_2.

*Existe una aparente ambigüedad en el signo D' al despejar su valor de la igualdad $\nu = D'^2\lambda_2$. El hecho de que se opte por un signo o por otro en el valor de D' se refleja solamente en la *dirección de positividad* del eje x'' en el sistema transformado (el sistema $x''y''$ proveniente del sistema original xy después de realizar en éste la rotación y traslación correspondiente). Como el objetivo en este ejercicio es solamente establecer la ecuación canónica de la parábola en cuestión, no existe ningún problema en que se opte por un signo o por otro para D'.

Los valores propios de la matriz M se obtienen al resolver la ecuación cuadrática

$$p(\lambda) = det\begin{bmatrix} A-\lambda & B \\ B & C-\lambda \end{bmatrix} = 0$$

Como se sabe, esta ecuación tiene sus dos raíces reales, pues M es una matriz simétrica. Sean λ_1 y λ_2 tales raíces.

Sea $u_1 = (x_1, y_1)$. Las coordenadas x_1 y y_1 (múltiplos de ellas) de este vector se obtienen al considerar alguna solución no trivial del sistema homogéneo

$$\begin{bmatrix} A-\lambda_1 & B \\ B & C-\lambda_1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Es decir que, en particular

$$(A-\lambda_1)x_1 + By_1 = 0$$

de donde

$$\frac{y_1}{x_1} = \frac{\lambda_1 - A}{B}^*$$

Obsérvese que el primer miembro de esta igualdad es la *pendiente de la recta* en la que se encuentra el vector u_1. Esta recta será el nuevo eje x del sistema girado (es decir, el eje x').

Similarmente se puede ver que

$$\frac{\lambda_2 - A}{B}$$

es la pendiente de la recta en la que se encuentra el vector u_2: el nuevo eje y del sistema girado (el eje y').

Resumiendo: al referir la ecuación (3.1) al nuevo sistema coordenado en el cual el nuevo eje x es la recta

$$y = \frac{\lambda_1 - A}{B}x \quad \text{(eje } x'\text{)}$$

y el nuevo eje y es la recta

$$y = \frac{\lambda_2 - A}{B}x \quad \text{(eje } y'\text{)}$$

*Se está suponiendo que $B \neq 0$, pues en caso contrario no habría necesidad de efectuar la simplificación por rotación de ejes coordenados en la ecuación (3.1).

tal ecuación se simplifica quedando de la forma (3.2)

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2D'x' + 2E'y' + F' = 0$$

Estúdiese el caso en el que $\mu \neq 0$ [caso en el que el lugar geométrico que representa la ecuación (3.1) es una curva con centro].

En este caso se ha visto que la ecuación canónica de la curva que representa la ecuación (3.1) es:

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \frac{\nu}{\mu} = 0$$

Esta ecuación carece de términos de primer grado. Partiendo de la ecuación (3.1) se puede lograr esta simplificación en ella (la eliminación de los términos de primer grado) si se realizan las sustituciones

$$x = x' + h$$
$$y = y' + k$$

y se escogen h y k convenientemente.

En efecto, al sustituir $x = x' + h$ y $y = y' + k$ en la ecuación (3.1) y simplificar se obtiene

$$Ax'^2 + 2Bx'y' + Cy'^2 + 2(Ah + Bk + D)x' + 2(Bh + Ck + E)y' + Ah^2 + 2Bhk + Ck^2 + 2Dh + 2Ek + F = 0$$

Entonces, si se escogen h y k tales que

$$Ah + Bk + D = 0$$
$$Bh + Ck + E = 0$$

se logrará una ecuación (simplificada) que carece de términos de primer grado.

Obsérvese que el sistema de ecuaciones anterior tiene una única solución (h, k) pues se está considerando el caso en el que μ (que es el determinante de este sistema) es no nulo.

Desde el punto de vista geométrico lo que se ha hecho es referir la ecuación (3.1) a un nuevo sistema coordenado cuyo origen está en el punto (h, k) (llamado el *centro* de la curva que representa tal ecuación). Ésta es pues la simplificación de la ecuación (3.1) por traslación de ejes coordenados.

Si se juntan estos resultados con aquellos provenientes de la simplificación de la ecuación (3.1) por rotación de ejes coordenados, se obtiene finalmente para el caso $\mu \neq 0$ que:

Al referir la ecuación

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

al nuevo sistema coordenado $x''y''$ en el cual el nuevo eje x es la recta

$$y - k = \frac{\lambda_1 - A}{B}(x - h) \quad (\text{eje } x'')$$

y el nuevo eje y es la recta

$$y - k = \frac{\lambda_2 - A}{B}(x - h) \quad (\text{eje } y'')$$

en donde h y k es la (única) solución del sistema

$$Ah + Bk + D = 0$$
$$Bh + Ck + E = 0$$

y λ_1 y λ_2 son los valores propios de la matriz

$$M = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

se obtiene la ecuación canónica de la curva que representa esta ecuación, la cual es:

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \frac{\nu}{\mu} = 0$$

Véanse algunos ejemplos.

EJEMPLO 9 Considere la ecuación

$$5x^2 + 6xy + 5y^2 - 4x + 4y - 4 = 0 \qquad (3.19)$$

Se tiene que

$$\mu = det\begin{bmatrix} A & B \\ B & C \end{bmatrix} = det\begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} = 16 > 0$$

$$\nu = det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = det\begin{bmatrix} 5 & 3 & -2 \\ 3 & 5 & 2 \\ -2 & 2 & -4 \end{bmatrix} = -128 \neq 0$$

$$\omega = A + C = 5 + 5 = 10$$

$$\omega\nu = (10)(-128) = -1280 < 0$$

Se trata entonces de una elipse.

Al resolver el sistema

$$Ah + Bk + D = 5h + 3k - 2 = 0$$
$$Bh + Ck + E = 3h + 5k + 2 = 0$$

se obtiene $h = 1$, $k = -1$.

Los valores propios de la matriz

$$M = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$$

son las raíces del polinomio

$$p(\lambda) = det\begin{bmatrix} 5-\lambda & 3 \\ 3 & 5-\lambda \end{bmatrix} = \lambda^2 - 10\lambda + 16 = (\lambda - 2)(\lambda - 8)$$

es decir, $\lambda_1 = 2$ y $\lambda_2 = 8$.

Al considerar entonces el nuevo eje x en la recta

$$y - k = \frac{\lambda_1 - A}{B}(x - h)$$

es decir,

$$y = -x \quad (\text{eje } x'')$$

y el nuevo eje y en la recta

$$y - k = \frac{\lambda_2 - A}{B}(x - h)$$

o sea,

$$y = x - 2 \quad (\text{eje } y'')$$

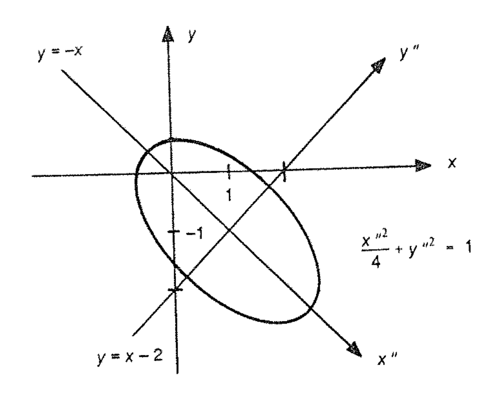
se tiene que la ecuación (3.19) se ve como

$$2x''^2 + 8y''^2 + \frac{-128}{16} = 0$$

o sea,

$$\frac{x''^2}{4} + y''^2 = 1$$

Gráficamente se tiene la siguiente situación:



EJEMPLO 10 Estudie ahora la ecuación

$$x^2 - 10xy + y^2 + x + y + 1 = 0 \qquad (3.20)$$

En este caso se tiene

$$\mu = det\begin{bmatrix} A & B \\ B & C \end{bmatrix} = det\begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix} = -24 < 0$$

$$\nu = det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = det\begin{bmatrix} 1 & -5 & 1/2 \\ -5 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{bmatrix} = -27 \neq 0$$

Se trata entonces de una hipérbola.

Al resolver el sistema

$$Ah + Bk + D = h - 5k + 1/2 = 0$$

$$Bh + Ck + E = -5h + k + 1/2 = 0$$

Se obtiene $h = k = 1/8$.

Los valores propios de la matriz

$$M = \begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix}$$

se obtienen hallando las raíces del polinomio

$$p(\lambda) = det\begin{bmatrix} 1-\lambda & -5 \\ -5 & 1-\lambda \end{bmatrix} = \lambda^2 - 2\lambda - 24 = (\lambda + 4)(\lambda - 6)$$

Éstos son: $\lambda_1 = -4$ y $\lambda_2 = 6$.

Al colocar entonces el nuevo eje x en la recta

$$y - \frac{1}{8} = \frac{-4-1}{-5}\left(x - \frac{1}{8}\right)$$

es decir,

$$y = x \quad (\text{eje } x'')$$

y el nuevo eje y en la recta

$$y - \frac{1}{8} = \frac{6-1}{-5}\left(x - \frac{1}{8}\right)$$

es decir,

$$y = -x + \frac{1}{4} \quad (\text{eje } y'')$$

se tiene que la ecuación (3.20) se ve como

$$-4x''^2 + 6y''^2 + \frac{-27}{-24} = 0$$

o sea,

$$\frac{x''^2}{9/32} - \frac{y''^2}{9/48} = 1$$

En la siguiente página se muestra la gráfica que resume los resultados de este ejemplo.*

EJEMPLO 11 Véase por último un ejemplo en el que la gráfica que representa la ecuación es una parábola. Sea

$$x^2 + 2xy + y^2 - \frac{10}{\sqrt{2}}x + \frac{2}{\sqrt{2}}y + 14 = 0 \qquad (3.21)$$

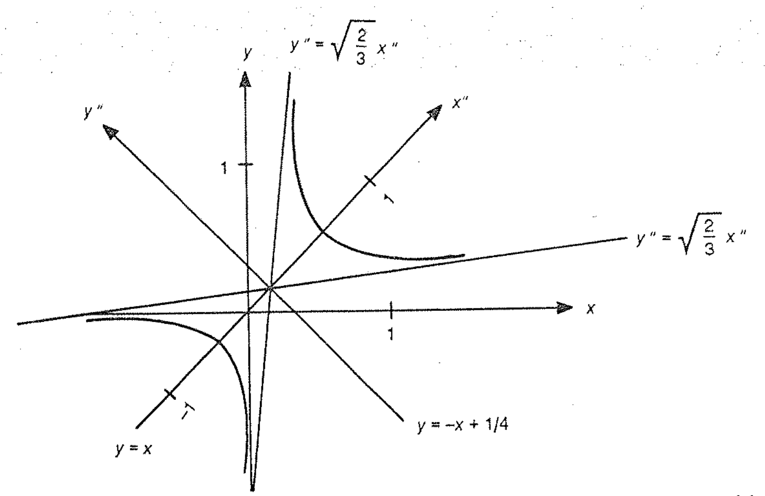
En este caso se tiene

$$\mu = det\begin{bmatrix} A & B \\ B & C \end{bmatrix} = det\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = 0$$

$$\nu = det\begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = det\begin{bmatrix} 1 & 1 & -5/\sqrt{2} \\ 1 & 1 & 1/\sqrt{2} \\ -5/\sqrt{2} & 1/\sqrt{2} & 14 \end{bmatrix} = -18 \neq 0$$

Se trata entonces de una parábola.

*Puede ser útil construir las gráficas de las asíntotas de esta hipérbola. Recuérdese que para la hipérbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ las rectas $y = \pm \frac{b}{a} x$ son las asíntotas de ella. En este caso se tiene que las rectas $y'' = \pm \sqrt{\frac{2}{3}}\, x''$ son las asíntotas (en el sistema $x''y''$) de la hipérbola en cuestión.



En este caso no se puede proceder como en los ejemplos anteriores, a establecer directamente las ecuaciones de los ejes x'' y y'' respecto de los cuales la ecuación (3.21) toma su forma canónica, pues aquí no es posible determinar el punto (h, k) a donde hay que trasladar el origen de coordenadas usando el sistema de ecuaciones

$$Ah + Bk + D = 0$$
$$Bh + Ck + E = 0$$

pues el determinante de este sistema es cero ($\mu = 0$).

Ya se ha visto que obtener la ecuación canónica de la parábola es un procedimiento fácil que no está comprometido con los valores de h y k mencionados anteriormente. Para esto sólo se necesita el valor propio λ_2 y el coeficiente D'.

Para la matriz

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

se tiene

$$p(\lambda) = det\begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} = \lambda(\lambda - 2)$$

de donde $\lambda_1 = 0$, $\lambda_2 = 2$.

Además se sabe que $\nu = -D'^2\lambda_2$, de donde

$$D' = \sqrt{-\frac{\nu}{\lambda_2}} = 3$$

Entonces la ecuación canónica de la parábola será:

$$\lambda_2 y''^2 + 2D'x'' = 0$$

es decir,

$$2y''^2 + 6x'' = 0$$

o bien,

$$y''^2 + 3x'' = 0 \tag{3.22}$$

Para construir la gráfica de la ecuación (3.21) se puede proceder de la manera "estándar" —como al principio de esta sección— al hacer primeramente la transformación por rotación de ejes y luego la traslación.

Se puede ver fácilmente que los vectores propios de M

$$u_1 = (1/\sqrt{2}, -1/\sqrt{2}) \text{ y } u_2 = (1/\sqrt{2}, 1/\sqrt{2})$$

forman una base ortonormal de $\mathbf{R}^2$. La matriz P de cambio de esta base canónica es entonces

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Al hacer entonces

$$\begin{bmatrix} x \\ y \end{bmatrix} = P\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} \tag{3.23}$$

o sea,

$$x = \frac{1}{\sqrt{2}}(x' + y')$$

$$y = \frac{1}{\sqrt{2}}(-x' + y')$$

y al sustituir en la ecuación (3.21) se obtiene, después de simplificar

$$y'^2 - 2y' - 3x' + 7 = 0 \tag{3.24}$$

Ésta es la ecuación (3.21) cuando se refiere al nuevo sistema coordenado $x'y'$ (girado respecto del sistema xy) en el cual el nuevo eje x es la recta

$$y = \frac{\lambda_1 - A}{B}x = -x \quad (\text{eje } x')$$

y el nuevo eje y es la recta

$$y = \frac{\lambda_2 - A}{B}x = x \quad (\text{eje } y')$$

Aún más, si se escribe la ecuación (3.24) como

$$(y' - 1)^2 - 3(x' - 2) = 0$$

y se hace

$$\begin{aligned} y'' &= y' - 1 \\ x'' &= x' - 2 \end{aligned} \qquad (3.25)$$

queda

$$y''^2 - 3x'' = 0$$

que es la misma ecuación canónica que ya se había obtenido (ec. 3.22) sólo que con el eje x'' orientado en la otra dirección (ver nota al pie de la página 659).

Las sustituciones (3.25) equivalen a trasladar el origen de coordenadas del sistema $x'y'$ al punto (2,1).

Si se quiere conocer las coordenadas de este punto en el sistema original se puede usar (3.23)

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

Entonces, con la notación que se ha venido usando para el nuevo origen del sistema coordenado, se tiene $h = 3/\sqrt{2}$, $k = -1/\sqrt{2}$. En conclusión, si se coloca el nuevo eje x en la recta

$$y + \frac{1}{\sqrt{2}} = -x + \frac{3}{\sqrt{2}}$$

o sea,

$$y = -x + \frac{2}{\sqrt{2}} \quad \text{(eje } x''\text{)}$$

y el nuevo eje y en la recta

$$y + \frac{1}{\sqrt{2}} = x - \frac{3}{\sqrt{2}}$$

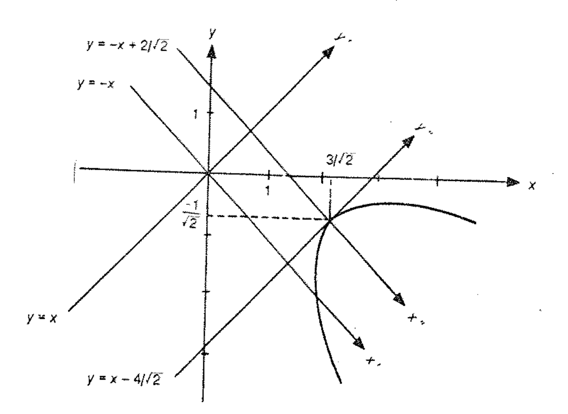
o sea,

$$y = x - \frac{4}{\sqrt{2}} \quad \text{(eje } y''\text{)}$$

la ecuación (3.21) se verá como

$$y''^2 - 3x'' = 0$$

Gráficamente:



NOTA: Otra manera de proceder para encontrar los valores de h y k es como sigue:

Se sabe de antemano que el nuevo eje y'' será *tangente* a la parábola representada por la ecuación (3.21) en el punto (h, k) —el vértice de la parábola—. Entonces en ese punto la derivada de la función $y = f(x)$ —implícita— en la ecuación (3.21) debe ser 1 = pendiente del eje y' = pendiente del eje y''.

Al derivar implícitamente (3.21) y al despejar y' se obtiene

$$y' = \frac{5/\sqrt{2} - x - y}{x + y + 1/\sqrt{2}}$$

entonces, cuando $x = h$ y $y = k$ debe cumplirse

$$\frac{5/\sqrt{2} - h - k}{h + k + 1/\sqrt{2}} = 1$$

Por otra parte, el punto (h, k) debe satisfacer la ecuación (3.21) pues él es vértice de la parábola. Es decir, que

$$h^2 + 2hk + k^2 - \frac{10}{\sqrt{2}}h + \frac{2}{\sqrt{2}}k + 14 = 0$$

Al resolver simultáneamente las dos últimas ecuaciones se obtiene que $h = 3/\sqrt{2}$ y $k = -1/\sqrt{2}$. Con estos valores, y los de las pendientes $\frac{\lambda_1 - A}{B}$ y $\frac{\lambda_2 - A}{B}$ de los ejes x'' y y'', se pueden establecer ya sus ecuaciones y concluir el problema.

EJERCICIOS (SECCIÓN 3, CAPÍTULO 7)

Cada una de las ecuaciones en los ejercicios del 1 al 15 representa una parábola, una elipse o una hipérbola. Simplifique la ecuación a su forma canónica e identifique el tipo de curva que ésta representa. Construya una gráfica de esta curva mostrando los sistemas coordenados obtenidos al simplificar la ecuación por rotación y por traslación de ejes.

1. $9x^2 - 4xy + 6y^2 + 16x - 8y - 2 = 0$
2. $x^2 - 2xy + y^2 - 10x - 6y + 25 = 0$
3. $2x^2 + 4xy + 5y^2 - 6x - 8y - 1 = 0$
4. $9x^2 - 24xy + 16y^2 - 20x + 110y - 50 = 0$
5. $\frac{1}{4}x^2 - \frac{1}{9}y^2 - x + \frac{2}{3}y - 1 = 0$
6. $4x^2 + 3y^2 + 16x + 12y - 36 = 0$
7. $3x^2 - 2xy + 3y^2 + 2\sqrt{2}x - 6\sqrt{2}y + 2 = 0$
8. $5x^2 + 12xy - 22x - 12y - 19 = 0$
9. $x^2 + 2xy + y^2 - 8x + 4 = 0$
10. $x^2 - 4xy + 4y^2 - 4x - 3y - 7 = 0$
11. $4x^2 - 12xy + 9y^2 - 8\sqrt{13}x - 14\sqrt{3}y + 117 = 0$
12. $5x^2 + 4xy + 8y^2 + 8x + 14y + 5 = 0$
13. $14x^2 + 24xy + 21y^2 - 4x + 18y - 139 = 0$
14. $5x^2 + 6xy + 5y^2 - 16x - 16y - 16 = 0$
15. $7x^2 + 16xy - 23y^2 - 14x - 16y - 218 = 0$
16. Demuestre que la ecuación

$$4x^2 - 4xy + y^2 - 6x + 3y - 4 = 0$$

representa dos líneas paralelas. Determine las ecuaciones de estas líneas en el sistema girado y en el sistema original.

17. Compruebe que la ecuación

$$3x^2 - 4xy - 4y^2 + 8x + 8y - 3 = 0$$

representa dos rectas que se cortan. Determine las ecuaciones de estas rectas.

18. Demuestre que la ecuación

$$5x^2 + 2xy + 10y^2 - 12x - 22y + 17 = 0$$

representa un punto en el plano. ¿Cuál es ese punto?

19. Demuestre que la ecuación

$$4x^2 - 20xy + 25y^2 + 4x - 10y + 1 = 0$$

representa una recta. Determine su ecuación.

20. Determine la naturaleza del lugar geométrico que representan las ecuaciones siguientes:
   a) $8x^2 - 18xy + 9y^2 + 2x - 1 = 0$
   b) $x^2 + 8xy + 16y^2 - 4x + 16y + 7 = 0$
   c) $x^2 + 2xy + y^2 + 2x + 2y + 1 = 0$
   d) $x^2 + 2xy + y^2 - 1 = 0$

21. Determine el tipo de lugar geométrico que representa la ecuación

$$x^2 + 2\alpha xy + y^2 - 1 = 0$$

para los diferentes valores de α (reales).

4. PARABOLOIDES, ELIPSOIDES, HIPERBOLOIDES, ETC.

En esta sección se estudiarán los lugares geométricos en $\mathbf{R}^3$* (llamados *superficies*) que representa la ecuación

$$Ax + By^2 + Cz^2 + 2Dxy + 2Exz + 2Fyz + 2Gx + 2Hy + 2Iz + J = 0 \quad (4.1)$$

El análisis que se hace para llegar a establecer una clasificación de estos lugares geométricos (similar a la presentada en la página 656 para la ecuación general de segundo grado) es exactamente el mismo que se presenta en la sección anterior (sólo que en este caso "con una coordenada más") de modo que se limitará a dar los resultados importantes de tal análisis para el caso de la ecuación (4.1).

Considérese la forma cuadrática $q: \mathbf{R}^3 \to \mathbf{R}$

$$q(x, y, z) = Ax^2 + By^2 + Cz^2 + 2Dxy + 2Exz + 2Fyz$$

cuya matriz respecto de la base canónica es:

$$M = \begin{bmatrix} A & D & E \\ D & B & F \\ E & F & C \end{bmatrix}$$

Se sabe que existe una base ortonormal β' de $\mathbf{R}^3$ respecto de la cual la forma q se escribe como

$$q(x, y, z) = \lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2$$

en donde $(x, y, z)_{\beta'} = (x', y', z')$ y $\lambda_1, \lambda_2, \lambda_3$ son los valores propios de la matriz M.

Sea P la matriz de cambio de la base de β' a la base canónica de $\mathbf{R}^3$.

Escríbase

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

Al sustituir en la ecuación (4.1) se obtiene entonces una expresión del tipo

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + 2G'x' + 2H'y' + 2I'z' + J' = 0 \quad (4.2)$$

*Se considera $\mathbf{R}^3$ como un espacio vectorial con el producto interno canónico.

Al igual que en la sección anterior se puede ver que al pasar de la ecuación (4.1) a la ecuación (simplificada) (4.2) lo que se hizo fue referir la ecuación original a un nuevo sistema coordenado $x'y'z'$ que se obtiene del sistema original xyz por medio de la rotación de éste un cierto ángulo θ. A los ejes x', y', z' se les llama *ejes principales* de la superficie que representa la ecuación (4.1).

Para lograr una nueva simplificación en la ecuación (4.2) —por traslación de ejes coordenados— se tendrán que considerar tres casos.

CASO 1. $\lambda_1\lambda_2\lambda_3 \neq 0$ (SUPERFICIE CON CENTRO).
En este caso se puede escribir la ecuación (4.2) como

$$\lambda_1\left(x'^2 + \frac{2G'}{\lambda_1}x'\right) + \lambda_2\left(y'^2 + \frac{2H'}{\lambda_2}y'\right) + \lambda_3\left(z'^2 + \frac{2I'}{\lambda_3}z'\right) + J' = 0$$

o bien, completando cuadrados

$$\lambda_1\left(x' + \frac{G'}{\lambda_1}\right)^2 + \lambda_2\left(y' + \frac{H'}{\lambda_2}\right)^2 + \lambda_3\left(z' + \frac{I'}{\lambda_3}\right)^2 + J' - \frac{G'^2}{\lambda_1} - \frac{H'^2}{\lambda_2} - \frac{I'^2}{\lambda_3} = 0$$

Haga

$$x'' = x' + G'/\lambda_1$$
$$y'' = y' + H'/\lambda_2$$
$$z'' = z' + I'/\lambda_3$$
$$J'' = J' - G'^2/\lambda_1 - H'^2/\lambda_2 - I'^2/\lambda_3$$

Entonces queda

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \lambda_3 z''^2 + J'' = 0 \tag{4.3}$$

Se deben considerar las dos posibilidades $J'' \neq 0$ y $J'' = 0$ por separado.

1.1. $J'' \neq 0$
Se puede suponer que J'' es un número negativo (caso contrario se cambia el signo en toda la ecuación), diga que $J'' = -j^2$. Entonces la ecuación (4.3) se puede presentar como

$$\frac{x''^2}{j^2/\lambda_1} + \frac{y''^2}{j^2/\lambda_2} + \frac{z''^2}{j^2/\lambda_3} = 1 \tag{4.4}$$

Se tienen cuatro posibilidades.

1.1.1. $\lambda_1 > 0,\ \lambda_2 > 0,\ \lambda_3 > 0$
Los números j^2/λ_i, $i = 1, 2, 3$ que aparecen en la ecuación (4.4) son en este caso positivos, diga que $j^2/\lambda_1 = a^2$, $j^2/\lambda_2 = b^2$, $j^2/\lambda_3 = c^2$, en donde a, b y c son números no nulos. Nos queda entonces la ecuación (4.4) como

$$\frac{x''^2}{a^2} + \frac{y''^2}{b^2} + \frac{z''^2}{c^2} = 1$$

que es la ecuación de un *elipsoide.*

1.1.2. $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 < 0$
En este caso la ecuación (4.4) puede escribirse como

$$\frac{x''^2}{a^2} + \frac{y''^2}{b^2} - \frac{z''^2}{c^2} = 1$$

para ciertos números no nulos a, b, c. Ésta es la ecuación de un *hiperboloide de una hoja*.

1.1.3. $\lambda_1 > 0,\ \lambda_2 < 0,\ \lambda_3 < 0$
En este caso la ecuación (4.4) puede escribirse como

$$\frac{x''^2}{a^2} - \frac{y''^2}{b^2} - \frac{z''^2}{c^2} = 1$$

para ciertos números no nulos a, b, c. Ésta es una ecuación de un *hiperboloide de dos hojas.*

1.1.4. $\lambda_1 < 0,\ \lambda_2 < 0, \lambda_3 < 0$
En este caso el primer miembro de la ecuación (4.4) es un número negativo. No existen entonces x'', y'', z'' que satisfagan tal ecuación. En este caso, pues, se tiene que la gráfica de la ecuación (4.4) —y por lo tanto la de la ecuación (4.1)— es un *conjunto vacío.*

1.2. $J'' = 0$
En este caso la ecuación (4.3) se ve como

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \lambda_3 z''^2 = 0 \tag{4.5}$$

Se tienen dos posibilidades.

1.2.1. TODOS LOS λ_i SON DE UN MISMO SIGNO.
La única solución para la ecuación (4.5) es en este caso $x'' = y'' = z'' = 0$. Ésta representa entonces *un punto.*

1.2.2. LOS λ_i TIENEN SIGNOS DISTINTOS.
Dígase por ejemplo que λ_1 y λ_2 son positivos y λ_3 negativo. Se puede entonces escribir la ecuación (4.5) como

$$z''^2 = a^2x''^2 + b^2y''^2$$

para ciertos números no nulos a y b. Ésta es la ecuación de *un cono.*

CASO 2. UNO DE LOS λ_i ES CERO.

Se puede suponer que $\lambda_3 = 0$ y $\lambda_1\lambda_2 \neq 0$. En este caso la ecuación (4.2) se ve como

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2G'x' + 2H'y' + 2I'z' + J' = 0$$

lo cual puede escribirse como

$$\lambda_1\left(x' + \frac{G'}{\lambda_1}\right)^2 + \lambda_2\left(y' + \frac{H'}{\lambda_2}\right)^2 + 2I'z' + J' - \frac{G'^2}{\lambda_1} - \frac{H'^2}{\lambda_2} = 0$$

o bien, haciendo

$$x'' = x' + G'/\lambda_1$$
$$y'' = y' + H'/\lambda_2$$
$$z'' = z'$$
$$J'' = J' - G'^2/\lambda_1 - H'^2/\lambda_2$$

queda como

$$\lambda_1 x''^2 + \lambda_2 y''^2 + 2I'z'' + J'' = 0 \qquad (4.6)$$

Las posibilidades $I' = 0$ e $I' \neq 0$ deben ser estudiadas por separado.

2.1. $I' = 0$

La ecuación (4.6) se ve en este caso como

$$\lambda_1 x''^2 + \lambda_2 y''^2 + J'' = 0 \qquad (4.7)$$

2.1.1. $J'' = 0$

La ecuación (4.7) se escribe como

$$\lambda_1 x''^2 + \lambda_2 y''^2 = 0 \qquad (4.8)$$

2.1.1.1. λ_1 Y λ_2 TIENEN EL MISMO SIGNO.

La única solución de la ecuación (4.8) es en este caso $x'' = y'' = 0$, que es la ecuación del eje z''. En este caso entonces la ecuación (4.6) —y por tanto la (4.1)— representa una *recta*.

2.1.1.2. λ_1 Y λ_2 TIENEN SIGNOS DISTINTOS.

Se puede escribir en este caso la ecuación (4.8) como

$$y''^2 = a^2 x''^2$$

o bien $y'' = \pm ax''$, en donde a es un cierto número no nulo. Éstas son las ecuaciones de *dos planos que se cortan* (en el eje z'').

2.1.2. $J'' \neq 0$

Se puede suponer que J'' es un número negativo, diga que $J'' = -j^2$, $j \neq 0$. En tal caso la ecuación (4.7) se ve como

$$\frac{x''^2}{j^2/\lambda_1} + \frac{y''^2}{j^2/\lambda_2} = 1 \qquad (4.9)$$

2.1.2.1. $\lambda_1 > 0$, $\lambda_2 > 0$

En este caso la ecuación (4.9) se ve como

$$\frac{x''^2}{a^2} + \frac{y''^2}{b^2} = 1$$

en donde a y b son ciertos números no nulos. Ésta es la ecuación de un *cilindro elíptico*.

2.1.2.2. $\lambda_1 > 0$, $\lambda_2 < 0$

En este caso la ecuación (4.9) se ve como

$$\frac{x''^2}{a^2} - \frac{y''^2}{b^2} = 1$$

en donde a y b son ciertos números no nulos. Ésta es la ecuación de un *cilindro hiperbólico*.

2.1.2.3. $\lambda_1 < 0$, $\lambda_2 < 0$

Como en este caso el primer miembro de la ecuación (4.9) es un número negativo, la gráfica que esta ecuación representa es un *conjunto vacío*.

2.2. $I' \neq 0$

Se puede escribir la ecuación (4.6) como

$$\lambda_1 x''^2 + \lambda_2 y''^2 + 2I'z'' = 0 \qquad (4.10)$$

en donde ahora

$$z'' = z' + J''/2I'$$

[x'' y y'' no cambian. Véanse ecuaciones que preceden a (4.6)].

2.2.1. λ_1 Y λ_2 TIENEN EL MISMO SIGNO

En este caso la ecuación (4.10) se puede escribir como

$$z'' = ax''^2 + by''^2$$

en donde a y b son ciertos números no nulos del mismo signo. Ésta es la ecuación de un *paraboloide elíptico*.

2.2.2. λ_1 Y λ_2 TIENEN SIGNOS CONTRARIOS

Se puede escribir en este caso la ecuación (4.10) como

$$z'' = ax''^2 - by''^2$$

en donde a y b son ciertos números no nulos del mismo signo. Ésta es la ecuación de un *paraboloide hiperbólico*.

CASO 3. DOS DE LOS λ_i SON CERO.

Se puede suponer que $\lambda_2 = \lambda_3 = 0$ y $\lambda_1 \neq 0$. En este caso la ecuación (4.2) se ve como

$$\lambda_1 x'^2 + 2G'x' + 2H'y' + 2I'z' + J' = 0$$

o bien,

$$\lambda_1 \left(x' + \frac{G'}{\lambda_1} \right)^2 + 2H'y' + 2I'z' + J'' = 0 \qquad (4.11)$$

en donde $J'' = J' - G'^2/\lambda_1$.

3.1. $H' = I' = 0$

Si hace

$$x'' = x' + G'/\lambda_1$$

$$y'' = y', \quad z'' = z'$$

la ecuación (4.11) queda escrita como

$$\lambda_1 x''^2 + J'' = 0 \qquad (4.12)$$

3.1.1. λ_1 Y J'' TIENEN EL MISMO SIGNO.

En tal caso es claro que no existe valor de x'' que satisfaga la ecuación (4.12). En este caso la gráfica de la ecuación (4.11) es un *conjunto vacío*.

3.1.2. λ_1 Y J'' TIENEN SIGNOS CONTRARIOS.

La ecuación (4.12) se ve en este caso como

$$x''^2 = a^2$$

para un cierto número no nulo a, o bien $x'' = \pm a$. Éstas son las ecuaciones de *dos planos paralelos* (al plano $y''z''$).

3.2. $H' \neq 0$ y/o $I' = 0$

Haga en este caso las siguientes sustituciones en la ecuación (4.11)

$$x'' = x' + G'/\lambda_1$$

$$y' = \frac{H'y'' + I'z''}{\sqrt{H'^2 + I'^2}} \qquad (4.13)$$

$$z' = \frac{I'y'' - H'z''}{\sqrt{H'^2 + I'^2}}$$

Después de simplificar, la ecuación (4.11) se transforma en

$$\lambda_1 x'''^2 + 2\sqrt{H'^2 + J'^2}\; y''' = 0$$

en donde $x''' = x''$ y $y''' = y'' + J''/2\,(H'^2 + I'^2)^{1/2}$, que es una ecuación del tipo

$$y''' = ax'''^2$$

para una cierta constante no nula a. Ésta es la ecuación de un *cilindro parabólico*.

En el cuadro de la siguiente página se presenta un resumen del análisis hecho anteriormente. También se presenta, después de este resumen, el aspecto gráfico que tienen algunas de las superficies que representa la ecuación (4.1).

Véanse un par de ejemplos.

EJEMPLO 1

Considere la ecuación

$$7x^2 + 6y^2 + 5z^2 - 4xy - 4yz - 6x - 24y + 18z + 18 = 0 \qquad (4.15)$$

La matriz M de la forma cuadrática q: $\mathbf{R}^3 \to \mathbf{R}$ dada por $q(x, y, z) = 7x^2 + 6y^2 + 5z^2 - 4xy - 4yz$, respecto de la base canónica de $\mathbf{R}^3$ es:

$$M = \begin{bmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{bmatrix}$$

cuyo polinomio característico es:

$$p(\lambda) = det\,(M - \lambda I) = det \begin{bmatrix} 7-\lambda & -2 & 0 \\ -2 & 6-\lambda & -2 \\ 0 & -2 & 5-\lambda \end{bmatrix} = -\lambda^3 + 18\lambda^2 - 99\lambda + 162$$

$$= -(\lambda - 3)(\lambda - 6)(\lambda - 9)$$

y entonces $\lambda_1 = 3$, $\lambda_2 = 6$, $\lambda_3 = 9$ son los valores propios de la matriz M (Caso 1).

Se pueden obtener fácilmente los vectores

$$u_1 = (1/3, 2/3, 2/3), \qquad u_2 = (2/3, 1/3, -2/3), \qquad u_3 = (2/3, -2/3, 1/3)$$

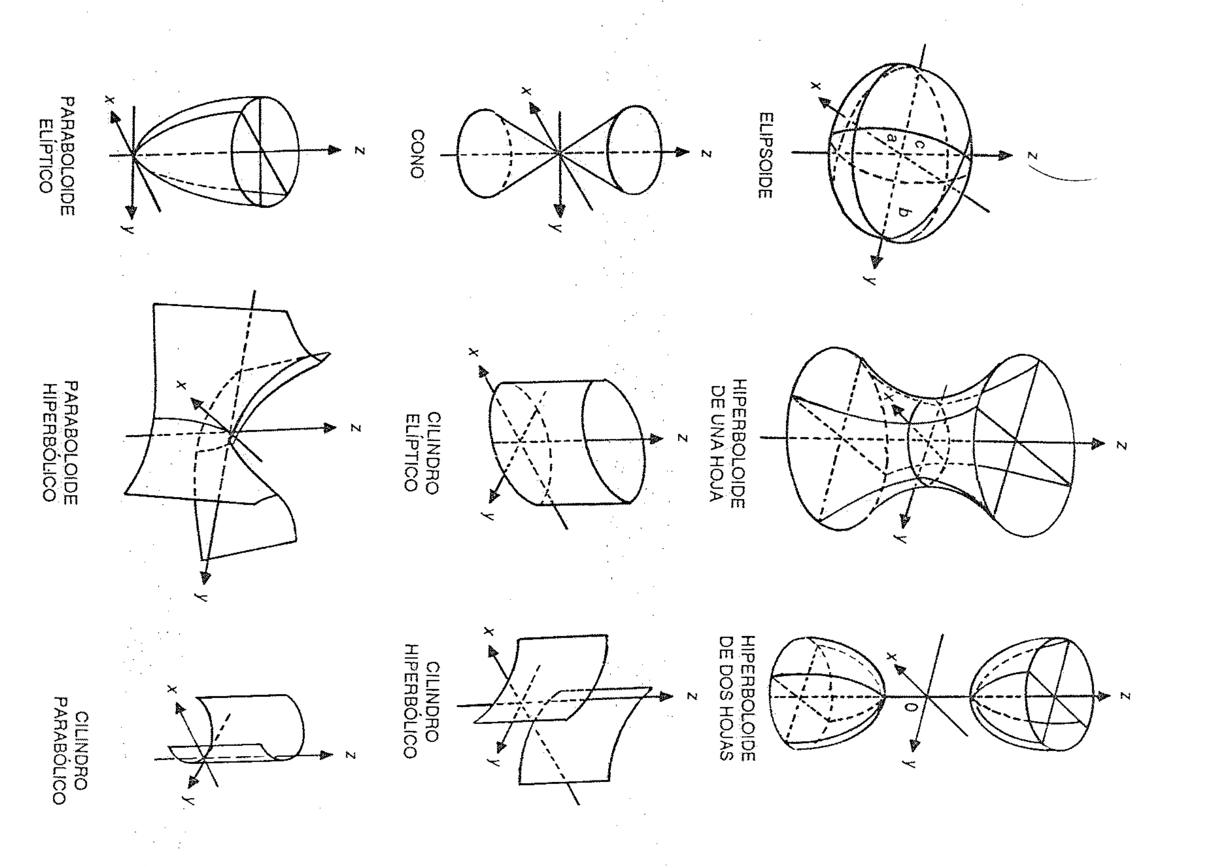
como vectores propios asociados a los valores propios λ_1, λ_2 y λ_3, respectivamente. Éstos forman una base ortonormal β' de $\mathbf{R}^3$.

ECUACIÓN: $Ax^2 + By^2 + Cz^2 + 2Dxy + 2Exz + 2Fyz + 2Gx + 2Hy + 2Iz + J = 0.$

λ_1, λ_2 y λ_3 son los valores propios de la matriz $M = \begin{bmatrix} A & D & E \\ D & B & F \\ E & F & C \end{bmatrix}$

LUGAR GEOMÉTRICO QUE REPRESENTA

- Caso 1. $\lambda_1\lambda_2\lambda_3 \neq 0$
  Ecuación simplificada
  $\lambda_1 x''^2 + \lambda_2 y''^2 + \lambda_3 z''^2 + J'' = 0$
  - $J'' \neq 0$ $(J'' < 0)$
    - $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0.$ **Un elipsoide.**
    - $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 < 0.$ **Un hiperboloide de una hoja.**
    - $\lambda_1 > 0, \lambda_2 < 0, \lambda_3 < 0.$ **Un hiperboloide de dos hojas.**
    - $\lambda_1 < 0, \lambda_2 < 0, \lambda_3 < 0.$ **Conjunto vacío.**
  - $J'' = 0$
    - $\lambda_1, \lambda_2, \lambda_3$ tienen el mismo signo. **Un punto.**
    - $\lambda_1, \lambda_2, \lambda_3$ tienen signos distintos. **Un cono.**
- Caso 2. $\lambda_3 = 0, \lambda_1\lambda_2 \neq 0$
  Ecuación simplificada
  $\lambda_1 x''^2 + \lambda_2 y''^2 + 2I'z'' + J'' = 0$
  - $I' = 0$
    - $J'' = 0$
      - λ_1, λ_2 tienen el mismo signo. **Una recta**
      - λ_1, λ_2 tienen signos contrarios. **Dos planos que se cortan**
    - $J'' \neq 0$
      - $\lambda_1 > 0, \lambda_2 > 0.$ **Un cilindro elíptico.**
      - $\lambda_1 > 0, \lambda_2 < 0.$ **Un cilindro hiperbólico.**
      - $\lambda_1 < 0, \lambda_2 < 0.$ **Conjunto vacío.**
  - $I'' \neq 0$
    - λ_1, λ_2 tienen el mismo signo. **Un paraboloide elíptico.**
    - λ_1, λ_2 tienen signos contrarios. **Un paraboloide hiperbólico.**
- Caso 3. $\lambda_2 = \lambda_3 = 0, \lambda_1 \neq 0$
  Ecuación simplificada
  $\lambda_1 x''^2 + 2H'y'' + 2'z'' + J'' = 0$
  - $H' = I' = 0$
    - λ_1, J'' tienen el mismo signo. **Conjunto vacío.**
    - λ_1, J'' tienen signos contrarios. **Dos planos paralelos.**
  - $H' \neq 0$ y/o $J' \neq 0.$ **Un cilindro parabólico**



La matriz P de cambio de base de β' a la base canónica es entonces

$$P = \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ 2/3 & 1/3 & -2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$$

Haciendo

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

o sea,

$$x = \frac{1}{3}(x' + 2y' + 2z')$$

$$y = \frac{1}{3}(2x' + y' - 2z')$$

$$z = \frac{1}{3}(2x' - 2y' + z')$$

y sustituyendo en la ecuación (4.15) se obtiene, después de simplificar

$$3x'^2 + 6y'^2 + 9z'^2 - 6x' - 24y' + 18z' + 18 = 0$$

o bien, completando cuadrados

$$3(x' - 1)^2 + 6(y' - 2)^2 + 9(z' + 1)^2 - 18 = 0$$

Al trasladar el origen de coordenadas al punto (1, 2, -1) (en el sistema $x'y'z'$) se obtiene finalmente que la ecuación (4.15) se escribe, en este nuevo sistema trasladado como

$$3x''^2 + 6y''^2 + 9z''^2 = 18$$

en donde $x'' = x' - 1$, $y'' = y' - 2$, $z'' = z' + 1$. Es decir,

$$\frac{x''^2}{6} + \frac{y''^2}{3} + \frac{z''^2}{2} = 0$$

que es la ecuación de un elipsoide.

EJEMPLO 2 Como último ejemplo considere la ecuación

$$2x^2 + 2y^2 + 3z^2 + 4xy + 2xz + 2yz - 4x + 6y - 2z + 3 = 0 \qquad (4.16)$$

La matriz M es en este caso

$$M = \begin{bmatrix} A & D & E \\ D & B & F \\ E & F & C \end{bmatrix} = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

Su polinomio característico es:

$$p(\lambda) = det(M - \lambda I) = det \begin{bmatrix} 2-\lambda & 2 & 1 \\ 2 & 2-\lambda & 1 \\ 1 & 1 & 3-\lambda \end{bmatrix} = -\lambda^3 + 7\lambda^2 - 10\lambda$$
$$= -\lambda(\lambda - 2)(\lambda - 5)$$

Los valores propios de M son entonces $\lambda_1 = 2$, $\lambda_2 = 5$, $\lambda_3 = 0$ (caso 2).

Una base ortonormal de $\mathbf{R}^3$ formada por vectores propios de M es:

$$\beta = \{(1/\sqrt{6}, 1/\sqrt{6}, -2/\sqrt{6}), (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}), (1/\sqrt{2}, -1/\sqrt{2}, 0)\}$$

Haciendo

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{6} & 1/\sqrt{3} & -1/\sqrt{2} \\ -2/\sqrt{6} & 1/\sqrt{3} & 0 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

y sustituyendo en la ecuación (4.16) se obtiene, después de simplificar

$$2x'^2 + 5y'^2 + \sqrt{6}x' - 5\sqrt{2}z' + 3 = 0$$

o bien, completando cuadrados

$$2(x' + \sqrt{6}/4)^2 + 5y'^2 - 5\sqrt{2}\,(z' - 9\sqrt{2}/40) = 0$$

Al hacer ahora $x'' = x' + \sqrt{6}/4$, $y'' = y'$, $z'' = z' - 9\sqrt{2}/40$ [esto es, trasladando el origen de coordenadas del sistema $x'y'z'$ al punto $(-\sqrt{6}/4, 0, 9\sqrt{2}/40)$] se obtiene

$$2x''^2 + 5y''^2 = 5\sqrt{2}z''$$

o bien,

$$z'' = \frac{\sqrt{2}}{5}x''^2 + \frac{\sqrt{2}}{2}y''^2$$

que es la ecuación de un paraboloide elíptico.

EJERCICIOS (SECCIÓN 4, CAPÍTULO 7)

En los ejercicios del 1 al 12 simplifique la ecuación a su forma canónica e identifique el tipo de lugar geométrico que ésta representa.

1. $2x^2 - 7y^2 - 4z^2 + 4xy + 20yz - 16xz + 60x - 12y + 12z - 90 = 0$
2. $2x^2 + 2y^2 - 5z^2 + 2xy - 2x - 4y - 4z + 2 = 0$
3. $3x^2 + 5y^2 + 3z^2 - 2xy + 2xz - 2yz - 12x - 10 = 0$
4. $2x^2 + 5y^2 + 2z^2 - 2xy - 4xz + 2yz + 2x - 10y - 2z - 1 = 0$
5. $x^2 - xy - xz + yz = 0$
6. $x^2 - 2y^2 + z^2 + 4xy + 4yz - 10xz + 2x + 4y - 10z + 17 = 0$
7. $2x^2 + y^2 + 2z^2 - 2xy - 2yz + x - 4y - 3z + 2 = 0$
8. $7x^2 + 6y^2 + 5z^2 - 4xy - 4yz - 6x - 24y + 18z + 30 = 0$
9. $4x^2 + y^2 - z^2 - 24x - 4y + 2z + 35 = 0$
10. $2x^2 + 2y^2 + 3z^2 + 4xy + 2yz + 2xz - 4x + 6y - 2z = 0$
11. $x^2 + 5y^2 + z^2 + 2xy + 6xz + 2yz - 6 = 0$
12. $x^2 + 4y^2 + 9z^2 + 12yz + 6xz + 4xy - 4x - 8y - 12z + 3 = 0$

④ 13. Al igual que para la ecuación general de segundo grado estudiada en la sección anterior (véase teorema 3.1), para la ecuación

$$Ax^2 + By^2 + Cz^2 + 2Dxy + 2Exz + 2Fyz + 2Gx + 2Hy + 2Iz + J = 0$$

se tienen ciertas cantidades *invariantes* cuando se realizan en ésta transformaciones por rotación y por traslación de ejes coordenados. Demuestre que si en esta ecuación se realizan las sustituciones

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = P \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

en donde P es una matriz ortogonal de orden 3, y/o las sustituciones

$$x = x' + h$$
$$y = y' + k$$
$$z = z' + \ell$$

en donde h, k y ℓ son ciertos números dados, entonces para la ecuación transformada

$$A'x'^2 + B'y'^2 + C'z'^2 + 2D'x'y' + 2E'x'z' + 2F'y'z' + 2G'x' + 2H'y' + 2I'z' + J' = 0$$

se conservan invariantes las cantidades

$$\omega = A + B + C$$

$$\mu = \det\begin{bmatrix} A & D \\ D & B \end{bmatrix} + \det\begin{bmatrix} A & E \\ E & C \end{bmatrix} + \det\begin{bmatrix} B & F \\ F & C \end{bmatrix}$$

$$\nu = \det\begin{bmatrix} A & D & E \\ D & B & F \\ E & F & C \end{bmatrix}$$

$$\rho = \det\begin{bmatrix} A & D & E & G \\ D & B & F & H \\ E & F & C & I \\ G & H & I & J \end{bmatrix}$$

(la idea de la demostración es exactamente la misma que la dada para el teorema 3.1). En particular demuestre que la ecuación (4.1) representa una superficie con centro (caso 1: $\lambda_1 \lambda_2 \lambda_3 \neq 0$) si, y sólo si $\nu \neq 0$. Use este hecho para demostrar que la ecuación canónica de una superficie con centro es:

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \lambda_3 z''^2 + \frac{\rho}{\nu} = 0$$

en donde λ_1, λ_2, λ_3 son los valores propios de la matriz

$$M = \begin{bmatrix} A & D & E \\ D & B & F \\ E & F & C \end{bmatrix}$$

(*Sugerencia*: use la ecuación 4.3 y la invariancia de ρ y ν.)

14. Use el ejercicio anterior para demostrar que

a) la ecuación

$$7x^2 + 6y^2 + 5z^2 - 4xy - 4yz - 6x - 24y + 18z + 30 = 0$$

representa un elipsoide cuya ecuación canónica es:

$$\frac{x''^2}{2} + y''^2 + \frac{x''^2}{2/3} = 1$$

b) la ecuación

$$x^2 + 5y^2 + z^2 + 2xy + 2yz + 6xz - 2x + 6y + 2z = 0$$

representa un hiperboloide de una hoja cuya ecuación canónica es:

$$\frac{x''^2}{1/3} + \frac{y''^2}{1/6} - \frac{z''^2}{1/2} = 1$$

c) la ecuación

$$2x^2 + 2y^2 - 5z^2 + 2xy - 2x - 4y - 4z + 2 = 0$$

representa un hiperboloide de dos hojas cuya ecuación canónica es:

$$\frac{x''^2}{4/5} + \frac{y''^2}{4/15} - \frac{z''^2}{4/25} = -1$$

d) la ecuación

$$x^2 + y^2 + z^2 + 2x + 4y - 6z + 15 = 0$$

no representa superficie alguna.

e) la ecuación

$$2x^2 + 3y^2 + z^2 + 12x - 12y + 8z + 46 = 0$$

representa un punto.

f) la ecuación

$$2x^2 + 5y^2 + 11z^2 + 20xy - 4xz + 16yz + 28x + 14y - 10z + 26 = 0$$

representa un cono cuya ecuación canónica es:

$$z''^2 = 2x''^2 + y''^2$$

Respuestas y soluciones a ejercicios seleccionados

INTRODUCCIÓN
(PÁGINAS 13-18)

1) a) $A \cup B = \{e, s, t, u, d, i, a, r, m, c, s\}$
$A \cup D = \{e, s, t, u, d, i, a, r, f, c, n\}$
$A \cap B = A \cap D = \{e, s, t, i, a\}$

b) $A \cap D = \{e, s, t, i, a\} \subset \{e, s, t, i, a, c\}$

d) $A \setminus D = A \setminus B = \{u, d, r\}$

2) a) $a \in B$ es correcto
b) $a \in A$ es correcto
c) $B \in A$ no se puede decidir
d) $\{a\} \in B$ es incorrecto
e) $\{a\} \subset B$ es correcto
f) $A \subset B$ no se puede decidir
g) $B \subset A$ es correcto
h) $B \subset \{a\}$ es correcto
i) $\{a\} \in A$ no se puede decidir
j) $\{a\} \subset A$ es correcto
k) $A \cap B = a$ es incorrecto
l) $a \in A \cup B$ es correcto
m) $\{a\} \subset A \cap B$ es correcto
n) $\{a\} \in A \cap B$ es incorrecto
o) $A \cup B = \{a\}$ no se puede decidir

3) a) $A \subset B$ es correcto
b) $A = B$ es correcto
c) $A \subset C$ es incorrecto
d) $A \subset D$ es incorrecto
e) $A = D$ es incorrecto
f) $E \subset F$ es correcto
g) $F \subset E$ es correcto
h) $A \in E$ es correcto
i) $A \subset E$ es incorrecto
j) $A \cap B = A$ es correcto
k) $E \cap F = \{1, 0\}$ es incorrecto
l) $E \cup F = \{0\} \cup A$ es correcto
m) $E \cap C = \{0\}$ es incorrecto
n) $E \cap D = \{0\}$ es correcto
o) $B \subset F$ es incorrecto

4) a) A es el conjunto vacío, es incorrecto
b) $\phi \subset A$ es correcto
c) $\phi \in A$ es correcto
d) $\{\phi\} \in A$ es correcto
e) $\{\phi\} \subset A$ es correcto
f) $A \cap \{\phi\} = \{\phi\}$ es correcto
g) $A \cap \phi = \phi$ es correcto
h) $A \cup \{\phi\} = A$ es correcto
i) $A \cap \{\phi\} = \phi$ es incorrecto
j) $A \cap \phi = \{\phi\}$ es incorrecto

5) Si $a = b$ el conjunto A es $A = \{a, a, \{a\}, \{a\}\} = \{a, \{a\}\}$. En este caso, es cierto que el conjunto A tiene sólo dos elementos distintos: a y $\{a\}$. Si $a \neq b$, el conjunto A tendría cuatro elementos distintos: a, $\{a\}$, b y $\{b\}$.

6) $\{0, 1, \{0\}, \{1\}, \{0, 1\}\}$

7) a) $p(A) = \{\phi\}$

b) $p(A) = \{\phi, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$

c) $p(p(A)) = \{\phi, \{\phi\}, \{\{1\}\}, \{\{2\}\}, \{\{1, 2\}\}, \{\phi, \{1\}\}, \{\phi, \{2\}\}, \{\phi, \{1, 2\}\},$
$\{\{1\}, \{2\}\}, \{\{1\}, \{1, 2\}\}, \{\{2\}, \{1, 2\}\}, \{\phi, \{1\}, \{2\}\},$
$\{\phi, \{1\}, \{1, 2\}\}, \{\phi, \{2\}, \{1, 2\}\}, \{\{1\}, \{2\}, \{1, 2\}\}, \{\phi, \{1\}, \{2\}, \{1, 2\}\}\}$

15) Una vez convencidos de que $\binom{n}{k}$ es el número de subconjuntos con k elementos de un conjunto A de n elementos, se ve que $p(A)$ tiene

$$\sum_{k=0}^{n} \binom{n}{k} = (1 + 1)^n = 2^n \text{ elementos}$$

17) Si $p(A) = p(B)$, entonces $A = B$, puesto que $x \in A \Rightarrow x \in p(A) \Rightarrow x \in p(B) \Rightarrow x \in B$.

22) $A \times \phi = \phi$

23) Si $A \times B = \phi$ entonces $A = \phi$ o $B = \phi$

24) b) De la inclusión $A \times C \subset B \times D$ *no* es posible concluir $A \subset B$ y $C \subset D$. Por ejemplo, ponga $C = \phi$ y A, B y D conjuntos arbitrarios.

25) a) Sí es una relación de equivalencia.

b) No es una relación de equivalencia (no es ni reflexiva, ni simétrica, ni transitiva).

c) No es una relación de equivalencia (no es necesariamente transitiva).

d) Sí es una relación de equivalencia.

e) Sí es una relación de equivalencia.

f) Sí es una relación de equivalencia.

29) $x \sim y \Leftrightarrow x - y = 5k$, algún $k \in z$

32) a) Rango de $f = \{1, 2, 3, 4\}$. La función sí es inyectiva. $f^{-1}: B \to A$ está dada por $\{(1, a), (2, b), (3, c), (4, d)\}$.

b) No es una función.

c) Rango de $f = \{3\}$. No es una función inyectiva.

d) Rango de $f = \{1, 2, 3, 4\}$. La función sí es inyectiva. $f^{-1}: B \to A$ está dada por $\{(1, a), (2, d), (3, c), (4, b)\}$.

e) Rango de $f = \{1, 2, 3\}$. No es una función inyectiva.

f) No es una función.

g) Rango de $f = \{1, 2, 3, 4\}$. La función sí es inyectiva. $f^{-1}: B \to A$ está dada por $\{(1, b), (2, c), (3, d), (4, a)\}$.

h) No es una función.

i) Rango de $f = \{2, 3\}$. No es una función inyectiva.

j) Rango de $f = \{1, 2\}$. No es una función inyectiva.

33) $a = 1, \quad b = 3, \quad c = -11 \quad d = -10$

34) a) Sí es lineal
b) No es lineal
c) No es lineal
d) No es lineal
e) Sí es lineal
f) No es lineal (a menos que $b = 0$).

35) Suponga que $f: A \to B$ es biyectiva. Entonces para cualquier $b \in B$ el conjunto $f^{-1}(b)$ no es vacío (pues f es sobreyectiva). Tome $x, x' \in f^{-1}(b)$. Entonces $f(x) = f(x') = b$, por lo que $x = x'$ (pues f es inyectiva). Es decir, $f^{-1}(b)$ consta de un solo elemento. Suponga ahora que la función $f: A \to B$ es tal que dado $b \in B$ el conjunto $f^{-1}(b)$ consta de un solo elemento. Entonces f es sobreyectiva y si $x, x' \in A$ son tales que $f(x) = f(x') = b$ se tiene que $x, x' \in f^{-1}(b)$ por lo que $x = x'$ y entonces, f es inyectiva. Es decir, f es biyectiva.

37) Son las 6 funciones $f_1, \ldots, f_6$: $\{1, 2, 3\} \to \{1, 2, 3\}$

$f_1(1) = 1$	$f_1(2) = 2$	$f_1(3) = 3$
$f_2(1) = 1$	$f_2(2) = 3$	$f_2(3) = 2$
$f_3(1) = 2$	$f_3(2) = 1$	$f_3(3) = 3$
$f_4(1) = 2$	$f_4(2) = 3$	$f_4(3) = 1$
$f_5(1) = 3$	$f_5(2) = 1$	$f_5(3) = 2$
$f_6(1) = 3$	$f_6(2) = 2$	$f_6(3) = 1$

38) $n!$

40) a) Suponga que $g \circ f: A \to C$ es inyectiva. En este caso, lo único que se puede decir es que f es una función inyectiva (demostración: sean $x_1, x_2 \in A$ tales que $f(x_1) = f(x_2)$. Entonces, $g(f(x_1)) = g(f(x_2))$, esto es, $(g \circ f)(x_1) = (g \circ f)(x_2)$ lo que implica que $x_1 = x_2$. La función g puede no ser inyectiva (por ejemplo: $f: \mathbb{R}^2 \to \mathbb{R}^3, f(x, y) = (x, y, z)$ y $g: \mathbb{R}^3 \to \mathbb{R}^2, g(x, y, z) = (x + z, y)$).

b) Suponga que $g \circ f: A \to C$ es sobreyectiva. En este caso, lo único que se puede decir es que g es una función sobreyectiva (demostración: dado $c \in C$ existe $a \in A$, tal que $(g \circ f)(a) = C$, esto es, dado $c \in C$ existe $f(a) \in B$ tal que $g(f(a)) = c$). La función f puede no ser sobreyectiva (considere las mismas funciones dadas como ejemplo al final del inciso anterior).

c) Suponga que $g \circ f: A \to C$ es biyectiva. Como consecuencia de los dos incisos anteriores, en este caso sólo se puede decir que f es una función inyectiva y que g es una función sobreyectiva.

41) Es una consecuencia directa de la definición de inversibilidad dada en la página 14.

42) Se tiene

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ \mathrm{id}_B \circ f = f^{-1} \circ f = \mathrm{id}_A$$

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ \mathrm{id}_B \circ g^{-1} = g \circ g^{-1} = \mathrm{id}_C$$

Entonces, $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$

44) Por inducción. Para $n = 1$ es parte de la hipótesis, esto es, $f(1) = 2(1) = 2$. Suponga el resultado válido para una $x \in \mathbb{N}$. Se tiene $f(x + 1) = f(x) + f(1) = 2x + 2 = 2(x + 1)$, lo que muestra que el resultado es válido para $x + 1$.

CAPÍTULO UNO

SECCIÓN 1 (PÁGINAS 28-30)

1) a) Sí es lineal
b) Sí es lineal
c) Sí es lineal
d) No es solución
e) No es lineal
f) Sí es lineal

2) a) No es solución
b) No es solución
c) Sí es solución
d) Sí es solución
e) Sí es solución
f) Sí es solución

3) El conjunto solución es vacío.

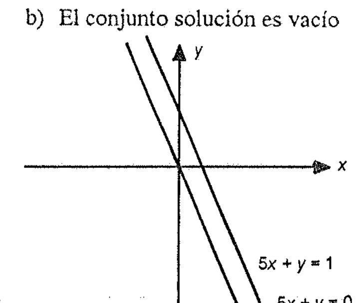
4) $\{x = 2\}$

5) a) $\{x, y \in \mathbf{R}\}$
b) Vacío
c) $\{x \in \mathbf{R}, y = 0\}$
d) $\{x \in \mathbf{R}, y = c/b\}$
e) $\{x = 0, y \in \mathbf{R}\}$
f) $\{x = c/a, y \in \mathbf{R}\}$
g) $\{x = t, y = -at/b, t \in \mathbf{R}\}$
h) $\{x = t, y = c - at/b, t \in \mathbf{R}\}$

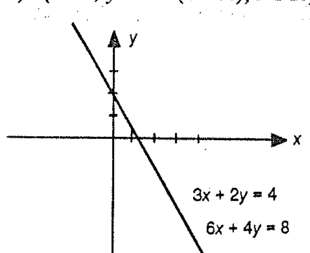
6) a) $\{x = 8/5, y = -12/5\}$

x − y = 4
(8/5, −12/5)
3x + 2y = 0

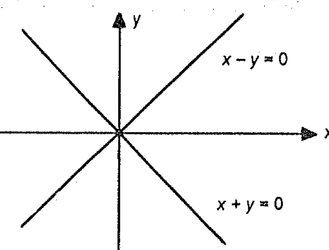
b) El conjunto solución es vacío



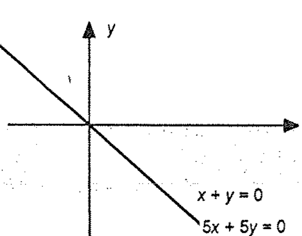
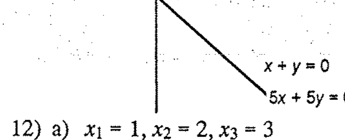
c) $\{x = t, y = 0.5\,(4 - 3t), t \in \mathbf{R}\}$



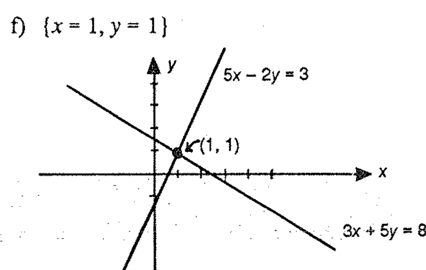
d) $\{x = 0, y = 0\}$



e) $\{x = t, y = -t, t \in \mathbf{R}\}$

f) $\{x = 1, y = 1\}$



12) a) $x_1 = 1, x_2 = 2, x_3 = 3$
b) $x_1 = -\dfrac{25}{31}, x_2 = -\dfrac{11}{31}, x_3 = \dfrac{43}{31}$
c) $x_1 = 1, x_2 = 1, x_3 = 2, x_4 = 2$

SECCIÓN 2 (PÁGINAS 50-54)

1) a) Matriz del sistema = $\begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}$

Matriz aumentada del sistema = $\begin{bmatrix} 2 & 1 & 0 \\ 3 & -1 & 5 \end{bmatrix}$

b) Matriz del sistema = $\begin{bmatrix} 2 & 3 & -4 \\ 3 & -4 & 7 \\ 6 & -7 & 8 \\ 8 & -4 & 7 \end{bmatrix}$

Matriz aumentada del sistema = $\begin{bmatrix} 2 & 3 & -4 & 7 \\ 3 & -4 & 7 & 4 \\ 6 & -7 & 8 & 4 \\ 8 & -4 & 7 & 9 \end{bmatrix}$

c) Matriz del sistema = $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Matriz aumentada del sistema = $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

d) Matriz del sistema = $\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$

Matriz aumentada del sistema = $\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$

2) a) $x_1 + 7x_2 + 8x_3 = 0$
$2x_1 + x_2 + 4x_3 = -1$

b) $x_1 + x_2 + x_3 = 1$
$x_1 + x_2 + x_3 = 1$
$x_1 + x_2 + x_3 = 1$

c) $x_1 + x_3 = 2$
$x_2 + x_4 = 3$
$x_1 + x_3 = 4$
$x_2 + x_4 = 5$

d) $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 4$
$x_2 = 2$
$x_1 + x_4 = 0$
$x_2 + x_3 + x_4 = 3$
$x_1 = 3$

e) $x_1 = 0$
$x_2 = 0$
$x_3 = 0$
$x_4 = 0$
$x_5 = 0$

3) a) Forma escalonada reducida.
b) Forma escalonada.
c) Forma escalonada reducida.
d) Forma escalonada reducida.
e) Forma escalonada reducida.

4) a) $\{x_1 = -3t, x_2 = 1 - 2t, x_3 = t, t \in \mathbf{R}\}$
b) $\{x_1 = 1 + 5t, x_2 = 1 - 3t, x_3 = t, t \in \mathbf{R}\}$
c) $\{x_1 = -2t, x_2 = -3t, x_3 = t, t \in \mathbf{R}\}$
d) $\{x_1 = 0, x_2 = r, x_3 = 2 - 3s, x_4 = 4 - 2t - s, x_5 = t, x_6 = s, r, s, t, \in \mathbf{R}\}$

5) a) Se intercambiaron las líneas 1 y 2 en A_1. Si se hace la misma operación en A_2, se recupera A_1.
b) Se sustituyó la tercera línea de A_1 por ella misma más 3 veces su primera línea. Si se sustituye la tercera línea en A_2 por ella misma menos 3 veces su primera línea se recupera A_1.
c) Se multiplicó la primera línea de A_1 por 7. Si se multiplica la primera línea de A_2 por 1/7 se recupera así A_1.

6) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & \alpha \\ 0 & 0 \end{bmatrix}$

7) Si, y sólo si $ad - bc \neq 0$.

9) Sean (L_i) y (L_j) las líneas i-ésima y j-ésima de la matriz A, respectivamente, realice los siguientes pasos:

1) $(L_i) \to (L_i) + (L_j)$
2) $(L_j) \to (L_j) + (L_i)$
3) $(L_i) \to (L_i) - (L_j)$
4) $(L_j) \to (L_j) + 2(L_i)$
5) $(L_i) \to (-1)(L_i)$

(cada paso está efectuado en la matriz inmediata anterior.)

11) a) $x_1 = \frac{47}{17}, x_2 = -\frac{3}{17}$
b) $x_1 = 1, x_2 = -2$
c) $x_1 = 1, x_2 = 0, x_3 = -1$
d) $x_1 = 2, x_2 = 2, x_3 = -2$
e) $x_1 = 0, x_2 = 0, x_3 = 1, x_4 = -1$
f) $x_1 = -32 + 72t, x_2 = 20 - 43t, x_3 = -7 + 17t, x_4 = t, t \in \mathbf{R}$
g) El sistema es inconsistente.
h) $x_1 = -54 + 87t - 79w$, $x_3 = -13 + 21t - 18w$
$x_2 = 32 - 53t + 47w$, $x_4 = t, x_5 = w, t, w \in \mathbf{R}$
i) $x_1 = -\frac{27}{44}t + \frac{31}{44}w$, $x_3 = \frac{7}{44}t + \frac{5}{44}w$
$x_2 = \frac{59}{44}t - \frac{71}{44}w$, $x_4 = t, x_5 = w, t, w \in \mathbf{R}$
j) $x_1 = -64 - 21t - 24w$, $x_3 = -15 - 4t - 5w$
$x_2 = 39 + 11t + 15w$, $x_4 = t, x_5 = w, t, w \in \mathbf{R}$

k) $x_1 = x_2 = x_3 = x_4 = x_5 = 1$
l) $x_1 = x_2 = x_3 = x_4 = 0$
m) $x_1 = t, x_2 = \frac{3}{4} - \frac{17}{3}t, x_3 = \frac{1}{3}t, x_4 = 1 + \frac{10}{3}t, x_5 = \frac{3}{4}, t \in \mathbf{R}$
n) $x_1 = x_2 = x_3 = 0, x_4 = x_5 = t, t \in \mathbf{R}$
o) $x_1 = -16 + t + w + 5v, x_2 = 23 - 2t - 2w - 6v, x_3 = t, x_4 = w, x_5 = v, t, w, v, \in \mathbf{R}$
p) $x_1 = 1, x_2 = 0, x_3 = 1, x_4 = 0, x_5 = 1, x_6 = 0$

12) La solución del sistema es $x_1 = 2 + x_3 + 2x_4 + \ldots + (n-2)x_n, x_2 = -1 - 2x_3 - 3x_4 - \ldots - (n-1)x_n$, en donde $x_3, x_4, \ldots, x_n$ son variables libres.

13) a) $a = \frac{4}{5}, b = \frac{17}{5}$
b) $a = -3, b = 5$
c) $a = 2, b = 1, c = 1$
d) $a = 1, b = 0, c = 0, d = 0$

14) Si $a \neq \pm 1$, la solución es $x_1 = \frac{2-a}{1-a^2}, x_2 = \frac{1-2a}{1-a^2}$

15) a) El sistema tiene una única solución, si $a \neq 0$ y $a \neq -\frac{1}{12}$. La solución es $x_1 = \frac{1}{a}$, $x_2 = 0$.
b) El sistema no tiene soluciones si $a = 0$.

16) $a = \pm 3$:
- $b \neq 2$. No hay solución.
- $b = 2$. Infinidad de soluciones. $(x1 = 1 - 13t, x_2 = 5t, x_3 = t, t \in \mathbf{R})$

$a \neq \pm 3$. Solución única.
$\left(x_1 = \frac{5a^2 - 2a^2b + 5b - 19}{a^2 - 9}, x_2 = \frac{(b-2)(a^2-4)}{a^2-9}, x_3 = \frac{b-2}{a^2-9}\right)$

17) $x_1 = \frac{2-7a}{4(a^2+2)}, x_2 = \frac{1}{2}, x_3 = \frac{a+7}{2(a^2+2)}$.

18) $\delta = 2$:
- $2a - 3b + c \neq 0$. No hay solución.
- $2a - 3b + c = 0$. Infinidad de soluciones. $(x_1 = b - a - t, x_2 = t, x_3 = 2a - b, t \in \mathbf{R}$

$\delta \neq 2$:
- $\delta = 1$:
  - $2a - 3b + c \neq 0$. No hay soluciones.
  - $2a - 3b + c = 0$. Infinidad de soluciones. $(x_1 = 2a - b - t, x_2 = b - a, x_3 = t, t \in \mathbf{R})$
- $\delta \neq 1$. Solución única.
  $\left(x_1 = \frac{2a - 3b + c}{(1-\delta)(2-\delta}, x_2 = \frac{c - b + a\delta - b\delta}{2-\delta}, x_3 = \frac{c - 2b + 2a\delta - b\delta}{\delta - 1}\right)$

19) 1) El sistema tiene una única solución si $a \neq 8$. Esta solución es:

$$x_1 = \frac{35a - b - 3ab - 80}{11(a-8)},\ x_2 = \frac{b(2a-3) - 5a - 64}{11(a-8)},\ x_3 = \frac{b-8}{a-8}$$

El sistema no tiene solución si $A = 8$ y $b \neq 8$.

El sistema tiene una infinidad de soluciones si $a = 8$ y $b = 8$. Estas soluciones son $x_1, = 1 - \frac{25}{11}t, x_2 = 1 + \frac{13}{11}t, x_3 = t, t \in \mathbf{R}$.

2) El sistema tiene una única solución si $a \neq 11$. Esta solución es:

$$x_1 = \frac{13b - a - 28}{a - 11},\ x_2 = \frac{a - 9b + 16}{a - 11},\ x_3 = \frac{b-3}{a-11}$$

El sistema no tiene solución si $a = 11$ y $b \neq 3$.

El sistema tiene una infinidad de soluciones si $a = 11$ y $b = 3$. Estas soluciones son: $x_1 = -1 + 13t, x_2 = 1 - 9t, x_3 = t, t \in \mathbf{R}$.

3) El sistema tiene una solución única si $a \neq 1$. Esta solución es:

$$x_1 = \frac{a-b-2}{a-1},\ x_2 = b - a + 3,\ x_3 = \frac{(2-a)(b-a+2)}{a-1}$$

El sistema no tiene solución si $a = 1$ y $b \neq -1$.

El sistema tiene una infinidad de soluciones si $a = 1$ y $b = -1$. Estas soluciones son: $x_1 = -t, x_2 = 1, x_3 = t, t \in \mathbf{R}$.

4) El sistema tiene una única solución si $a \neq \pm 2$. Esta solución es:

$$x_1 = \frac{1 - ab - 2a}{4 - a^2},\ x_2 = \frac{2b + 4 - a}{4 - a^2},\ x_3 = 0$$

El sistema no tiene solución si $a = 2$ y $b \neq 0$.

El sistema tiene una infinidad de soluciones si $a = 2$ y $b = 0$ (que son: $x_1 = \frac{1}{2} - t - 2w, x_2 = t, x_3 = w, t, w \in \mathbf{R}$) y también si $a = -2$ (que son $x_1 = t - \frac{1}{4}(b+1), x_2 = t, x_3 = \frac{1}{8}(b+3), t \in \mathbf{R}$).

20) 1) Si $a = \pm 1$, hay una infinidad de soluciones $x_1 = 1 - at, x_2 = t, t \in \mathbf{R}$.
Si $a \neq \pm 1$, la solución es única, $x_1 = 1, x_2 = 0$.

2) Si $a = 1$, hay una infinidad de soluciones $x_1 = \frac{1}{2} - t, x_2 = t, x_3 = \frac{1}{2}, t \in \mathbf{R}$. Si $a \neq 1$, la solución es única, $x_1 = 1, x_2 = -\frac{1}{4}, x_3 = \frac{1}{4}$.

3) Si $a = \pm 2$, hay una infinidad de soluciones $x_1 = \frac{5}{3} - \frac{2}{3}t, x_2 = \frac{4}{3} - \frac{1}{3}t, x_3 = t, t \in \mathbf{R}$.
(Si $a \neq \pm 2$ no hay soluciones.)

4) Si $a = \pm 3$ hay una infinidad de soluciones $x_1 = 5r + t - 14, x_2 = -2r, x_3 = r, x_4 = t, r, t \in \mathbf{R}$ (si $a \neq \pm 3$, no hay soluciones).

21) 1) $x_1 = -\frac{19}{10}t, x_2 = \frac{7}{10}t, x_3 = t, t \in \mathbf{R}$

2) $x_1 = x_2 = x_3 = 0$

3) $x_1 = x_2 = x_3 = 0$

4) $x_1 = \frac{1}{3}t, x_2 = -\frac{19}{15}t, x_3 = \frac{2}{5}t, x_4 = t, t \in \mathbf{R}$

5) $x_1 = -\frac{2}{3}t, x_2 = \frac{5}{3}t, x_3 = t, t \in \mathbf{R}$

22) Es falso para sistemas no homogéneos. Por ejemplo, $x_1 + x_2 + x_3 = 1, x_1 + x_2 + x = 2$.

SECCIÓN 3 (PÁGINAS 73-80)

1) $a = \frac{7}{3}, b = -\frac{8}{3}$

2) $a = 27, b = 17$

3) $a = -1, b = 2, c = -3, d = 4$

12) a) Todas las matrices cuadradas de orden n.

b) Todas las matrices cuadradas de orden n.

c.1) $B = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}, a, b \in \mathbf{R}$ c.3) $B = \begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{bmatrix}, a, b, c \in \mathbf{R}$

c.2) $B = \begin{bmatrix} b & 3a \\ a & b \end{bmatrix}, a, b \in \mathbf{R}$

19) c.1) $[18 \ \ 7 \ \ 4]$ c.2) $\begin{bmatrix} 7 \\ 12 \\ 9 \end{bmatrix}$ c.3) $\begin{bmatrix} 4 \\ 13 \\ 9 \end{bmatrix}$

20) a) $tr\,A = 15, tr\,B = 11, tr\,A + tr\,B = 26$

b) $tr\,(A + B) = 26$

c) $tr\,(AB) = tr\,(BA) = 82$

24) $AB = \begin{bmatrix} 11 & 5 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 5 & 0 & 0 \\ 0 & 0 & 6 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 14 & 19 \\ 0 & 0 & 0 & 0 & 14 & 18 \end{bmatrix}$

27) a) $AB = \begin{bmatrix} 18 & 25 & 74 \\ 13 & 17 & 50 \\ 4 & 4 & -4 \end{bmatrix}$ c) $AB = \begin{bmatrix} 10 & 21 & 29 & 39 \\ 17 & 16 & 26 & 36 \\ 6 & 6 & 5 & 4 \\ 12 & 28 & 5 & 5 \end{bmatrix}$

b) $AB = \begin{bmatrix} 24 & 15 & 11 & 62 \\ 20 & 13 & 9 & 50 \\ 18 & 14 & 13 & 74 \\ 6 & 11 & 3 & 28 \end{bmatrix}$

28) $\begin{bmatrix} 4 & 5 & 1 & -2 & 1 & 2 & 0 & 0 & 0 & 0 \\ 10 & 7 & 4 & -1 & 1 & 3 & 0 & 0 & 0 & 0 \\ -2 & 3 & -1 & -4 & 1 & 2 & 0 & 0 & 0 & 0 \\ 3 & 1 & 2 & -4 & -5 & -6 & 0 & 0 & 0 & 0 \\ 4 & 5 & -1 & -3 & -1 & -24 & 0 & 0 & 0 & 0 \\ 3 & 1 & 1 & -4 & -2 & 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & 2 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 27 & 14 & 29 & -4 & -7 & 8 & 0 & 0 & 1 & 0 \\ 75 & 32 & 76 & -7 & -7 & -4 & 0 & 0 & 0 & 1 \end{bmatrix}$

29) a) $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1-t \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -t \\ t \end{bmatrix}$

b) $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 1/2 + t \\ t \end{bmatrix} = \begin{bmatrix} 3/2 \\ 1/2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ t \\ t \end{bmatrix}$

c) $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

d) $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{7}{5} - \frac{2}{5}t \\ \frac{2}{5} + \frac{3}{5}t \\ t \end{bmatrix} = \begin{bmatrix} \frac{7}{5} \\ \frac{2}{5} \\ 0 \end{bmatrix} + \begin{bmatrix} -\frac{2}{5}t \\ \frac{3}{5}t \\ t \end{bmatrix}$

30) a) La matriz es $\begin{bmatrix} 1 & 1 & 2 & 0 & 1 \\ 1 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1/2 \end{bmatrix}$

b) La matriz es $\begin{bmatrix} 20 & 20 & 20 & 20 & 30 & 30 \\ 30 & 30 & 30 & 30 & 40 & 40 \\ 10 & 10 & 10 & 10 & 20 & 20 \\ 20 & 20 & 20 & 20 & 40 & 40 \\ 30 & 30 & 30 & 30 & 40 & 40 \end{bmatrix}$

c) La matriz es $\begin{bmatrix} 100 & 100 & 100 & 100 & 150 & 150 \\ 40 & 40 & 40 & 40 & 60 & 60 \\ 20 & 20 & 20 & 20 & 30 & 30 \\ 25 & 25 & 25 & 25 & 40 & 40 \\ 25 & 25 & 25 & 25 & 40 & 40 \end{bmatrix}$

Se consumen 280 refrescos de cola/semana, 180 limonadas y 35 botellas de ron/semana.

SECCIÓN 4 (PÁGINAS 98-105)

7) $B^{-1} = \begin{bmatrix} 1 & 2/3 & 4/3 \\ 2/3 & 1/3 & -1/3 \\ 1/3 & 1/3 & 4/3 \end{bmatrix}$

10) Puesto que la matriz de coeficientes del sistema es inversible, la única solución de éste es: $x_1 = x_2 = x_3 = 0$.

13) De la expresión $A^3 - 3A^2 - 7A + 18I_3 = 0$, se obtiene que $A\left(\frac{1}{18}(7I_3 + 3A - A^2)\right) = I_3$. Entonces,

$$A^{-1} = \frac{1}{18}(7I_3 + 3A - A^2),\ A^{-1} = \begin{bmatrix} 1/6 & 1/3 & 1/2 \\ -11/198 & 5/9 & 1/6 \\ 1/9 & -1/9 & -1/3 \end{bmatrix}$$

14) $A^{-1} = \frac{1}{11}(A^2 - 3A - 7I_3),\ A^{-1} = \begin{bmatrix} -2/11 & 5/11 & -1/11 \\ -1/11 & -3/11 & 5/11 \\ 7/11 & -1/11 & -2/11 \end{bmatrix}$

15) a) No es elemental.

b) Sí es elemental. Su inversa es $\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$

c) No es elemental.

d) Sí es elemental. Su inversa es $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$

16) Sea $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

a) $E_1A = \begin{bmatrix} a & b & c \\ 3d & 3e & 3f \\ g & h & i \end{bmatrix}$

b) $E_2A = \begin{bmatrix} a+4g & b+4h & c+4i \\ d & e & f \\ g & h & i \end{bmatrix}$

c) $E_3A = \begin{bmatrix} a & b & c \\ g & h & i \\ d & e & f \end{bmatrix}$

d) $E_1E_3A = \begin{bmatrix} a & b & c \\ 3g & 3h & 3i \\ d & e & f \end{bmatrix}$

e) $E_1E_2A = \begin{bmatrix} a+4g & b+4h & c+4i \\ 3d & 3e & 3f \\ g & h & i \end{bmatrix}$

f) $E_2E_1A = E_1E_2A$ (inciso anterior)

g) $E_2E_3A = \begin{bmatrix} a+4d & b+4e & c+4f \\ g & h & i \\ d & e & f \end{bmatrix}$

h) $E_3E_1A = \begin{bmatrix} a & b & c \\ g & h & i \\ 3d & 3e & 3f \end{bmatrix}$

i) $E_3E_1E_2A = \begin{bmatrix} a+4g & b+4h & c+4i \\ g & h & i \\ 3d & 3e & 3f \end{bmatrix}$

j) $E_3^2E_2A = E_2A$ (inciso b)).

k) $E_2^{-1}E_3^2E_1A = \begin{bmatrix} a-4g & b-4h & c-4i \\ 3d & 3e & 3f \\ g & h & i \end{bmatrix}$

l) $E_1^{-1}E_2E_3E_2^{-1}A = \begin{bmatrix} c-4g+4d & b-4h+4e & c-4i+4f \\ g/3 & h/3 & i/3 \\ d & e & f \end{bmatrix}$

17) Si E es la matriz elemental correspondiente, se tiene $(EA)(B) = E(AB)$. Entonces, cualquier operación elemental que se haga en las líneas de A corresponderá con la misma operación elemental en las líneas de AB.

18) Si E es la matriz elemental que provino de I haciendo la operación elemental en las líneas de I-e, entonces AE es la matriz que proviene de A haciendo en ésta la operación elemental e *en las columnas de* A.

19) Si E es la matriz elemental correspondiente, se tiene $A(BE) = (AB)E$. Entonces, cualquier operación elemental que se haga en las columnas de B, corresponderá con la misma operación elemental hecha en las columnas de AB.

20) Sea E la matriz elemental que proviene de I haciendo en ésta la operación elemental e (inciso a), b) o c)).

Sea $E' = EA$. Entonces $(A')^{-1} = A^{-1}E^{-1}$. Entonces, si en A se hace la operación elemental e en sus líneas, en la inversa de A se efectuará la operación elemental inversa a e (la que corresponde a la matriz elemental E^{-1}) *en sus columnas*.

21) Si en A se hace la operación elemental e en sus columnas, en la inversa de A se efectuará la operación elemental inversa a e en sus líneas.

28) C es invertible puesto que con $C^{-1} = \begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix}$ se tiene $CC^{-1} = C^{-1}C = I$.

$$C^{-1} = \begin{bmatrix} 1 & -3 & 0 & 0 \\ -2 & 7 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

29) a) $A^{-1} = \begin{bmatrix} -1/5 & 3/5 \\ 2/5 & -1/5 \end{bmatrix}$

b) $A^{-1} = \begin{bmatrix} 3/4 & -1 \\ -1/2 & 1 \end{bmatrix}$

c) No es inversible

d) No es inversible

e) $A^{-1} = \begin{bmatrix} 27/57 & -4/19 & 27/57 \\ 1/19 & 8/19 & -2/19 \\ -2/19 & 3/19 & 4/19 \end{bmatrix}$

f) No es inversible

g) $A^{-1} = \begin{bmatrix} -1/3 & -7/3 & 2 \\ -1/3 & 5/3 & -1 \\ 1 & -2 & 1 \end{bmatrix}$

h) $A^{-1} = \begin{bmatrix} 7/39 & 1/13 & 22/39 & 9/13 \\ -1/39 & -2/13 & -31/39 & -5/13 \\ 8/39 & 3/13 & 14/39 & 1/13 \\ -5/39 & 3/13 & 1/39 & 1/13 \end{bmatrix}$

i) $A^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

j) $A^{-1} = \begin{bmatrix} 1 & -2 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 4 & -1 & -2 \\ 0 & -2 & 1 & 1 \end{bmatrix}$

k) $A^{-1} = \begin{bmatrix} -6 & -26 & 22 & 17 \\ 5 & 20 & -17 & -13 \\ 0 & 2 & -1 & -1 \\ -1 & -5 & 4 & 3 \end{bmatrix}$

l) $A^{-1} = \begin{bmatrix} 11/6 & -5/3 & 5/4 & -7/6 & 5/2 \\ 1/6 & 1/3 & 1/12 & 1/6 & -1/2 \\ -1/2 & 2/3 & -7/12 & 1/2 & -1/2 \\ 2/3 & -2/3 & 2/3 & -1/3 & 1 \\ 1/2 & -1/3 & 5/12 & -1/2 & 1/2 \end{bmatrix}$

m) $A^{-1} = \frac{1}{68}\begin{bmatrix} 20 & 2 & -14 & -14 & 6 & -22 \\ 79 & 115 & 11 & 130 & 22 & 78 \\ -30 & 14 & 38 & 4 & 8 & 16 \\ 13 & 105 & 13 & 98 & 26 & 86 \\ 19 & 41 & -15 & 36 & 4 & 8 \\ -6 & -38 & -6 & -40 & -12 & -24 \end{bmatrix}$

30) La matriz B^{-1} es: $B^{-1} = \begin{bmatrix} 25/32 & 3/32 & 1/16 & -1/2 \\ 3/16 & 1/16 & 3/8 & 0 \\ 5/8 & -1/8 & 1/4 & 0 \\ -33/32 & 5/32 & -9/16 & 1/2 \end{bmatrix}$

La matriz A es: $A = \begin{bmatrix} 8 & 1 & 26 & 18 & 22 & 5 & 5 & 20 \\ 21 & 22 & 4 & 9 & 1 & 19 & 13 & 22 \\ 3 & 8 & 16 & 1 & 12 & 7 & 5 & 2 \\ 19 & 1 & 12 & 9 & 14 & 5 & 1 & 12 \end{bmatrix}$

El mensaje es: HAY QUE ESTUDIAR MUCHO ÁLGEBRA LINEAL.

31) a) $X = \begin{bmatrix} 11/7 & 1/7 \\ 6/7 & 5/7 \end{bmatrix}$ c) $X = \begin{bmatrix} -7/2 & 19/10 \\ 2 & -4/5 \end{bmatrix}$

b) $X = \begin{bmatrix} 1 & 5 \\ -2 & -17 \end{bmatrix}$ d) $X = \begin{bmatrix} 91 & -53 \\ -36 & 21 \end{bmatrix}$

32) a) $X = \begin{bmatrix} 1 & 3 & 2 \\ -4/5 & -1 & 1/2 \\ 1/5 & -2 & -2 \end{bmatrix}$ c) $X = \begin{bmatrix} -56/11 & -31/11 & 41/11 \\ -47/11 & 27/11 & 35/11 \\ -56/11 & -31/11 & 41/11 \end{bmatrix}$

b) $X = \begin{bmatrix} 1/5 & 6/5 & 37/50 \\ 0 & 1 & 1 \\ -1/10 & -31/10 & -68/25 \end{bmatrix}$ d) $X = \begin{bmatrix} 1/2^3 & 0 & 0 \\ 0 & 1/(2)^4(3)^3 & 0 \\ 0 & 0 & 1/(3)^4(4)^3 \end{bmatrix}$

33) $A = E_1E_2E_3E_4E_5$, en donde

$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ $E_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1/3 \end{bmatrix}$

$E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ $E_5 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4/3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

$E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 4 & 1 \end{bmatrix}$

(la respuesta no es única.)

34) Sea E_i la matriz elemental que provino de I_n multiplicando la i-ésima línea de ésta por el escalar a_i, $i = 1, 2, \ldots, n$. Entonces, si $A = diag\,(a_1, a_2, \ldots, a_n)$, se tiene $A = E_1E_2 \ldots E_n$.

35) $A = (A) \quad (I)$

$\uparrow$ inversible $\quad \uparrow$ forma escalonada reducida de A

36) a) $\begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -10 \\ 0 & 1 & 4 \end{bmatrix}$

b) $\begin{bmatrix} 3 & 1 & 4 \\ 2 & 1 & 3 \\ -1 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 0 \\ 2 & 1 & 1 \\ -1 & -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

c) $\begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 0 \\ 1 & 4 & 0 \\ 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

37) a) $x_1 = 2, \quad x_2 = -1$

b) $x_1 = -1, \quad x_2 = -1/2, \quad x_3 = 5/4$

c) $x_1 = \frac{131}{85}, \quad x_2 = \frac{173}{85}, \quad x_3 = \frac{122}{85}, \quad x_4 = \frac{73}{85}$

CAPÍTULO DOS

SECCIÓN 1 (PÁGINAS 115-118)

1) Son las dos permutaciones (4, 3, 1, 2) y (4, 3, 2, 1)

2) Son las 6 permutaciones

(6, 5, 4, 1, 2, 3) (6, 5, 4, 2, 3, 1)
(6, 5, 4, 1, 3, 2) (6, 5, 4, 3, 1, 2)
(6, 5, 4, 2, 1, 3) (6, 5, 4, 3, 2, 1)

3) a) $\sigma \circ \pi = (3, 1, 2)$
$\pi \circ \sigma = (2, 3, 1)$
$\sigma \circ \sigma = (1, 2, 3)$
$\pi \circ \pi = (1, 2, 3)$

b) $\sigma \circ \pi = (3, 1, 4, 2)$
$\pi \circ \sigma = (4, 3, 1, 2)$
$\sigma \circ \sigma = (2, 4, 3, 1)$
$\pi \circ \pi = (1, 2, 3, 4)$

c) $\sigma \circ \pi = \pi \circ \sigma = \sigma \circ \sigma$
$= \pi \circ \pi = (1, 2, 3, 4)$

d) $\sigma \circ \pi = (2, 5, 4, 1, 3)$
$\pi \circ \sigma = (5, 1, 2, 3, 4)$
$\sigma \circ \sigma = (3, 2, 5, 4, 1)$
$\pi \circ \pi = (1, 2, 3, 4, 5)$

e) $\sigma \circ \pi = (3, 5, 4, 2, 1)$
$\pi \circ \sigma = (4, 5, 2, 3, 1)$
$\sigma \circ \sigma = (1, 2, 3, 4, 5)$
$\pi \circ \pi = (2, 1, 3, 5, 4)$

f) $\sigma \circ \pi = (2, 1, 7, 5, 4, 6, 3)$
$\pi \circ \sigma = (5, 2, 4, 3, 1, 7, 6)$
$\sigma \circ \sigma = (1, 2, 3, 4, 5, 6, 7)$
$\pi \circ \pi = (2, 5, 6, 1, 3, 7, 4)$

4) a) $\sigma^{-1} = (3, 2, 1)$
$\pi^{-1} = (1, 3, 2)$
$(\sigma \circ \pi)^{-1} = \pi^{-1} \circ \sigma^{-1} = (2, 3, 1)$

b) $\sigma^{-1} = (2, 4, 3, 1)$
$\pi^{-1} = (3, 2, 1, 4)$
$(\sigma \circ \pi)^{-1} = \pi^{-1} \circ \sigma^{-1} = (2, 4, 1, 3)$

c) $\sigma^{-1} = \pi^{-1} = (\sigma \circ \pi)^{-1} =$
$= \pi^{-1} \circ \sigma^{-1} = (1, 2, 3, 4)$

d) $\sigma^{-1} = (3, 2, 5, 4, 1)$
$\pi^{-1} = (2, 1, 4, 3, 5)$
$(\sigma \circ \pi)^{-1} = \pi^{-1} \circ \sigma^{-1} = (4, 1, 5, 3, 2)$

e) $\sigma^{-1} = (1, 2, 4, 3, 5)$
$\pi^{-1} = (5, 4, 3, 1, 2)$
$(\sigma \circ \pi)^{-1} = \pi^{-1} \circ \sigma^{-1} = (5, 4, 1, 3, 2)$

f) $\sigma^{-1} = (7, 6, 5, 4, 3, 2, 1)$
$\pi^{-1} = (3, 6, 4, 5, 7, 1, 2)$
$(\sigma \circ \pi)^{-1} = \pi^{-1} \circ \sigma^{-1} = (2, 1, 7, 5, 4, 6, 3)$

6)

Inversiones/signo

	a)	b)	c)	d)	e)	f)
σ	3 / -1	4 / +1	5 / -1	6 / +1	1 / -1	21 / -1
π	1 / -1	3 / -1	5 / -1	2 / +1	9 / -1	12 / +1
$\sigma \circ \pi$	2 / +1	3 / -1	0 / +1	6 / +1	8 / +1	9 / -1
$\pi \circ \sigma$	2 / +1	5 / -1	0 / +1	4 / +1	8 / +1	9 / -1
$\sigma \circ \sigma$	0 / +1	4 / +1	0 / +1	6 / +1	0 / +1	0 / +1
$\pi \circ \pi$	0 / +1	0 / +1	0 / +1	0 / +1	2 / +1	8 / +1

7)

Inversiones/signo

	a)	b)	c)	d)	e)	f)
σ^{-1}	3 / -1	4 / +1	5 / -1	6 / +1	1 / -1	21 / -1
π^{-1}	1 / -1	3 / -1	5 / -1	2 / +1	9 / -1	12 / +1
$(\sigma \circ \pi)^{-1}$	2 / +1	3 / -1	0 / +1	6 / +1	8 / +1	9 / -1

8) a) $\frac{n(n-1)}{2}$ inversiones

b) $sgn\ \sigma = (-1)^{\frac{n(n-1)}{2}}$

c) $\sigma^{-1} = \sigma$

9) a) 3 inversiones. $\sigma^{-1} = \sigma$

b) 3 inversiones $\sigma^{-1} = \sigma$

c) 9 inversiones. $\sigma^{-1} = \sigma$

d) 17 inversiones. $\sigma^{-1} = \sigma$

10) $\sigma^{-1} = \sigma$

11)

b.1) $A_\sigma = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ $\qquad A_\pi = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

b.2) $A_{\sigma \circ \pi} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

d) $A_\sigma = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \; A_{\sigma^{-1}} = A_\sigma$

SECCIÓN 2 (PÁGINAS 134-139)

1) a) Sí aparece con signo +.
b) Sí aparece con signo -.
c) No aparece (tiene iguales elementos en la 1a. columna)
d) Sí aparece con signo +.
e) Sí aparece con signo +.
f) No aparece (dos elementos en la 1a. línea).
g) No aparece (dos elementos en la 1a. columna).
h) No aparece (dos elementos en la 5a. columna).

2) $a_{11}a_{23}a_{32}a_{44}$ con signo -
$a_{12}a_{23}a_{31}a_{44}$ con signo +
$a_{14}a_{23}a_{31}a_{42}$ con signo -
$a_{11}a_{23}a_{34}a_{42}$ con signo +
$a_{12}a_{23}a_{34}a_{41}$ con signo -
$a_{14}a_{23}a_{32}a_{41}$ con signo +

3) $a_{11}a_{23}a_{34}a_{43}a_{55}$ con signo -
$a_{11}a_{23}a_{34}a_{42}a_{55}$ con signo +
$a_{11}a_{25}a_{34}a_{42}a_{53}$ con signo -
$a_{11}a_{22}a_{34}a_{45}a_{53}$ con signo +
$a_{11}a_{23}a_{34}a_{45}a_{52}$ con signo -
$a_{11}a_{25}a_{34}a_{43}a_{52}$ con signo +

4) a) $a_{11}a_{22}a_{33}a_{44}a_{55}a_{66}$
$a_{11}a_{22}a_{33}a_{45}a_{56}a_{64}$
$a_{11}a_{22}a_{33}a_{46}a_{54}a_{65}$
$a_{11}a_{22}a_{34}a_{43}a_{56}a_{65}$
$a_{11}a_{22}a_{34}a_{45}a_{53}a_{66}$
$a_{11}a_{22}a_{34}a_{46}a_{55}a_{63}$
$a_{11}a_{22}a_{35}a_{43}a_{54}a_{66}$
$a_{11}a_{22}a_{35}a_{46}a_{53}a_{64}$
$a_{11}a_{22}a_{35}a_{44}a_{56}a_{63}$
$a_{11}a_{22}a_{36}a_{43}a_{55}a_{64}$
$a_{11}a_{22}a_{36}a_{44}a_{53}a_{65}$
$a_{11}a_{22}a_{36}a_{45}a_{54}a_{63}$

b) $a_{11}a_{24}a_{32}a_{43}a_{55}a_{66}$
$a_{11}a_{25}a_{32}a_{43}a_{56}a_{64}$
$a_{11}a_{26}a_{32}a_{43}a_{54}a_{65}$

c) $a_{13}a_{21}a_{32}a_{44}a_{56}a_{65}$
$a_{13}a_{21}a_{32}a_{45}a_{54}a_{66}$
$a_{13}a_{21}a_{32}a_{46}a_{55}a_{64}$

d) Sólo el producto $a_{11}a_{22}a_{33}a_{44}a_{56}a_{65}$

5) $i = 3, j = 1$

6) Se tienen 3 posibilidades
$i = 7, j = 6, k = 5$
$i = 5, j = 7, k = 6$
$i = 6, j = 5, k = 7$

7) El coeficiente es -10.

8) El coeficiente es +18.

9) a) Sí aparece. $k = -12$
b) Sí aparece. $k = -6$
c) No aparece
d) No aparece

10) El determinante es $(-1)^{[[\frac{n}{2}]]} = (-1)^{\frac{n(n-1)}{2}}$ (en donde $[[x]]$ denota el mayor entero menor que x).

12) El determinante vale $(-1)^{\left[\frac{n}{2}\right]} a_{1n}a_{2,n-1}\dots a_{n1}$

13) a) -15　　c) -34
b) 38　　d) 0

17) a) -3　　d) 12
b) -18　　e) 3
c) -3　　f) 3

18) $det\ A_c = det\ A$

19) a) $x_1 = 2,\ x_2 = 3,\ x_3 = 4$
b) $x_1 = a,\ x_2 = b,\ x_3 = c,\ x_4 = -(a+b+c)$

20) Las n soluciones son: $x_1 = 0, x_2 = 1, \dots, x_n = n - 1$

21) a) 0　　d) 0
b) 0　　e) $(b-a)(c-a)(c-b)$
c) 0　　f) 0

23) a) 5　　e) 768
b) 0　　f) -3
c) 1　　g) 0
d) -24　　h) -52

SECCIÓN 3 (PÁGINAS 143-145)

3) Si A es la matriz: $A = \begin{bmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{bmatrix}$

Al multiplicar "líneas por líneas" (como en el inciso b) del ejercicio 1) se obtiene la matriz B dada por $B = (a^2 + b^2 + c^2 + d^2)\, I^4$.

Entonces, por los resultados del ejercicio 2, se tiene $det\ B = (a^2 + b^2 + c^2 + d^2)^4 = (det\ A)^2$ de donde $det\ A = (a^2 + b^2 + c^2 + d^2)^2$

4) $det\ A^2 = (det\ A)^2 \geq 0$

5) $det\ (c\,A) = det\,(cIA) = det\ (cI)(det\ A) = c^n\, det\ A$

6) a) A es inversible pues $det\ A = 2 \neq 0$　　d) $det\ (5A)^{-1} = \frac{1}{(5)^3(2)} = \frac{1}{250}$
b) $det\ A^{-1} = 1/2$　　e) $det\ (2A^{-1})^4 = 2^8$
c) $det\ 5A^{-1} = (5^3)\left(\frac{1}{2}\right) = \frac{125}{2}$

7) a) AB es inversible, pues, $det(AB) = (det\ A)(det\ B) = (3)(-2) = -6 \neq 0$.
La matriz BA también es inversible, pues, $det\ (BA) = -6 \neq 0$.
b) $det\ (AB)^{-1} = -1/6 = det\ (BA)^{-1}$
c) $det\ (A^{-1}B) = -2/3$
d) $det\ (2A^{-1})(3B^{-1}) = -216$
e) $det\ (5B^{-1})(2A)(6B) = 3.888 \times 10^7$
f) $det\ (5B)^{-1}(2A)(6B) = \frac{62208}{625}$

8) A_σ es inversible, pues, $A_\sigma = sgn\ \sigma \neq 0$. Se tiene $A_\sigma^{-1} = A_{\sigma^{-1}}$

11) a) $tr\,A = n(n+1)$

12) $det\ A = 0$

13) $det\ A = \pm 1$.

SECCIÓN 4 (PÁGINAS 156-160)

1) $C_{12} = -18,\ C_{31} = 7,\ C_{22} = 4,\ C_{33} = 7,\ C_{13} = -9$

2) -46

3) a) 0 (su segunda columna está formada por ceros)
b) $-70a$

4) Desarrolle el determinante por cofactores a lo largo de la línea o la columna de ceros.

5) a) 4　　e) 0
b) -6　　f) -22
c) 519　　g) -5
d) 519

7) a) $det\ A = 8$　　c) $det\ A_2 = det\ Q = 1$
b) $det\ A_1 = det\ P = 8$

9) $det\ A = -3$

10) $det\ A = -652$

SECCIÓN 5 (PÁGINAS 167-171)

1) a) $adj\ A = \begin{bmatrix} -10 & 5 & 5 \\ 10 & -4 & -3 \\ -5 & 3 & 1 \end{bmatrix}$　　$adj\ B = \begin{bmatrix} 14 & -10 & 22 \\ 44 & -13 & 20 \\ -18 & 19 & -16 \end{bmatrix}$

b) $(adj B)(adj A) = \begin{bmatrix} -350 & 176 & 122 \\ -670 & 332 & 279 \\ 450 & -214 & -163 \end{bmatrix}$

c) $adj\ AB = (adj B)(adj A)$

2) Cierto

3) $\begin{bmatrix} 63 & -27 & -11 \\ 0 & 9 & -8 \\ 0 & 0 & 7 \end{bmatrix}$

6) $adj\,A = (det A)\,A = \pm A$

7) En este caso A no es inversible. Entonces $A(adj\,A) = (det\,A)\,I = 0 \cdot I = 0$.

8) Si A es nilpotente, A no es inversible (véase ejercicio 25 de la sección 4, capítulo 1). Entonces $A(adj\,A) = (det\,A)\,I = 0 \cdot I = 0$.

9) De la fórmula $A(adj\,A) = (det\,A)I$ se tiene $det\,(A(adj\,A)) = det\,((det\,A)I)\,(det\,A)(det(adj\,A)) = (det\,A)^n\,det\,I$. Como $det\,A \neq 0$ y $det\,I = 1$, se obtiene finalmente, $det\,(adj\,A) = (det\,A)^{n-1}$.

12) De la fórmula $A(adj\,A) = (det\,A)I$ poniendo en lugar de A a su matriz adjunta se obtiene $(adj\,A)(adj(adj\,A)) = det\,(adj\,A)I$. Multiplicando por la derecha por A se obtiene $A(adj\,A)(adj(adj\,A)) = (det\,(adj\,A))A$ o sea $(det\,A)(adj(adj\,A)) = (det\,A)^{n-1}\,A$, de donde $adj(adj\,A) = (det\,A)^{n-2}\,A$ y entonces $det\,(adj(adj\,A)) = det\,((det\,A)^{n-2}\,A) = ((det\,A)^{n-2})^n\,det\,A = (det\,A)^{(n-1)^2}$

13) a) $\frac{1}{2}\begin{bmatrix} 2 & 0 & 0 \\ 2 & 4 & -2 \\ -2 & -1 & 1 \end{bmatrix}$ c) $-\frac{1}{32}\begin{bmatrix} -3 & -6 & -1 \\ -17 & -2 & 5 \\ 8 & 16 & -8 \end{bmatrix}$

b) $\frac{1}{69}\begin{bmatrix} 27 & -1 & -20 \\ -21 & 11 & 13 \\ 12 & -3 & 9 \end{bmatrix}$

14) Como $A^{-1} = \dfrac{1}{det\,A}\,adj\,A = adj\,A$ y todos los cofactores de A están formados por sumas y productos de enteros, se concluye que $A^{-1} = adj\,A$ tiene sólo elementos enteros.

16) En este caso la matriz A_i, $i = 1, 2, \ldots, n$ tiene su i-ésima columna formada exclusivamente de ceros. Entonces, $x_i = \dfrac{det\,A_i}{det\,A} = 0$, $i = 1, 2, \ldots, n$

17) a) $x_1 = 1/2\; x_2 = 0$ d) $x_1 = -13, x_2 = 0, x_3 = 20$

b) $x_1 = 37/60, x_2 = -29/60$ e) $x_1 = -10/9, x_2 = 57/27, x_3 = 12/27$

c) $x_1 = 113/5, x_2 = 11, x_3 = -42/5$

18) a.1) $AB = \begin{bmatrix} 34 & 11 \\ 12 & -1 \end{bmatrix}$

a.2) $det\,AB = -166$

a.3) $A(\ ,1) = \begin{bmatrix} 4 & 3 \\ 1 & 1 \end{bmatrix}$

$A(\ ,2) = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$

$A(\ ,3) = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$

$B(1,\ ) = \begin{bmatrix} 3 & 4 \\ 7 & -1 \end{bmatrix}$

$B(2,\ ) = \begin{bmatrix} 1 & -2 \\ 7 & -1 \end{bmatrix}$

$B(3,\ ) = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}$

a.4) $A(\ ,1)B(1,\ ) = \begin{bmatrix} 33 & 13 \\ 10 & 3 \end{bmatrix}$

$A(\ ,2)B(2,\ ) = \begin{bmatrix} 22 & -5 \\ 9 & -5 \end{bmatrix}$

$A(\ ,3)B(3,\ ) = \begin{bmatrix} 13 & 14 \\ 5 & 0 \end{bmatrix}$

d.1) $-166 = det\,AB(1,3) = det\begin{bmatrix} 34 & 11 \\ 12 & -1 \end{bmatrix} = det\,A(1,1)B(1,3) + det\,A(1,2)B(2,3)$

$+\,det\,A(1,3)B(3,3) = det\begin{bmatrix} 33 & 13 \\ 10 & 3 \end{bmatrix} + det\begin{bmatrix} 22 & -5 \\ 9 & -5 \end{bmatrix} + det\begin{bmatrix} 13 & 14 \\ 5 & 0 \end{bmatrix}$

$= -31 - 65 - 70$

d.2) $321 = det\,AB(2,1) = det\begin{bmatrix} 7 & 146 \\ -1 & 25 \end{bmatrix} = det\,A(2,1)B(1,1)$

$+\,det\,A(2,2)B(2,1) + det\,A(2,3)B(3,1) = det\begin{bmatrix} 27 & 66 \\ 3 & 9 \end{bmatrix}$

$+\,det\begin{bmatrix} -25 & 90 \\ -5 & 18 \end{bmatrix} + det\begin{bmatrix} 12 & 136 \\ 0 & 23 \end{bmatrix} = 45 + 0 + 276$

e.3) $-2066 = det\,AB(3,2) = det\begin{bmatrix} 69 & 146 \\ 29 & 42 \end{bmatrix} = det\,A(3,1)B(1,2)$

$+\,det\,A(3,2)B(2,2) + det\,A(3,3)B(3,2) = det\begin{bmatrix} 59 & 66 \\ 33 & 34 \end{bmatrix}$

$+\,det\begin{bmatrix} 45 & 90 \\ 22 & 14 \end{bmatrix} + det\begin{bmatrix} 34 & 136 \\ 13 & 36 \end{bmatrix} = -172 - 1350 - 544$

f) $adj\,AB = (det\,AB)(AB)^{-1} = (det\,A)(det\,B)B^{-1}A^{-1} = ((det\,B)B^{-1})((det\,A)A^{-1})$
$= (adj\,B)(adj\,A)$

APÉNDICE (PÁGINAS 174-175)

1) Es una consecuencia directa de la fórmula (A1) de la página 179.

2) a) $(3-2)(3-1)(2-1) = 2$

b) $(-2-2)(-2-(-1))(-2-1)(2-(-1))(2-1)(-1-1) = 72$

c) $(-3-3)(-3-1)(-3-0)(-3-(-1))(3-1)(3-0)(3-(-1))(1-0)(1-(-1))$
$(0-(-1)) = 6912$

3) Al imponer las condiciones $p(x_i) = y_i$, $i = 1, 2, 3$, se obtiene el sistema de tres ecuaciones lineales en las indeterminadas a, b, c

$$x_1^2a + x_1b + c = y_1,\quad x_2^2a + x_2b + c = y_2,\quad x_3^2a + x_3b + c = y_3$$

El determinante de este sistema es V_3. Como los números x_1, x_2, x_3 son distintos, se tiene que $V_3 \neq 0$. Entonces, el sistema tiene una única solución para a, b, c. O sea que existe el polinomio $p(x) = ax^2 + bx + c$ con las condiciones requeridas.

CAPÍTULO TRES

SECCIÓN 1 (PÁGINAS 182-183)

2) No es un espacio vectorial: la suma no es conmutativa ni asociativa.

3) a) Sí es un espacio vectorial.
   b) No es un espacio vectorial (no existe el cero en el conjunto).
   c) No es un espacio vectorial (no todo vector tiene inverso aditivo).
   d) No es un espacio vectorial (no se cumple $1 \cdot x = x$).

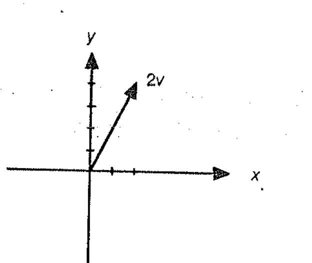
SECCIÓN 2 (PÁGINAS 195-198)

1) a) (0, 0, 0, 0)
   b) (−1, 2, 0, −3, −8)
   c) (4, 6, 2)
   d) (0, 0, 0, 0, 0, 0,0)
   e) (40, −5, 15, 10)
   f) (−4, −4, −4, −12, −32, 4, −28)
   g) (−15, −15, −21, −18, −12)
   h) (−2, 2, −16, −10, −20, −8, 0, −2)
   i) (0, 0, 0, 0, 0)
   j) (1, 1, 1, 1, 1)
   k) $(x_1, x_2, x_3, x_4) + (y_1, y_2, y_3, y_4) = (y_1, y_2, y_3, y_4) + (x_1, x_2, x_3, x_4)$
   l) (34, 6, 26, 2)
   m) (5, 7, 9, 10, 45, 86, −70)
   n) (−10, −30, 10, −40, −80)
   o) (7/3, 1/3, −1/3, 0, 4/3)

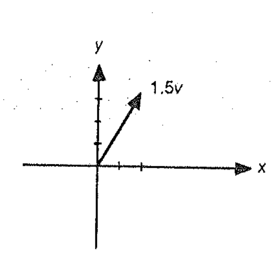
3) a) es equivalente a v_2
   b) no es equivalente a ninguno de ellos
   c) es equivalente a v_3
   d) es equivalente a v_1
   e) es equivalente a v_2
   f) es equivalente a v_1
   g) es equivalente a v_3
   h) no es equivalente a ninguno de ellos

5)

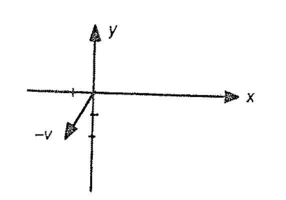
a)



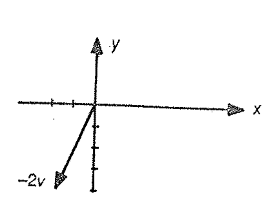
b)



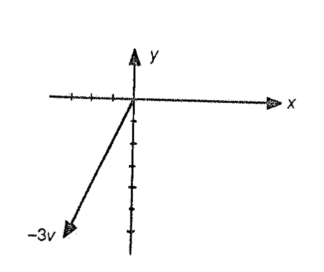
c)



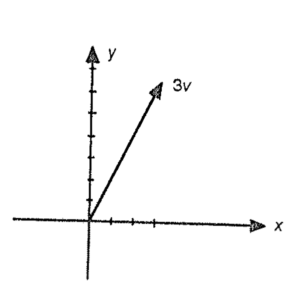
d)



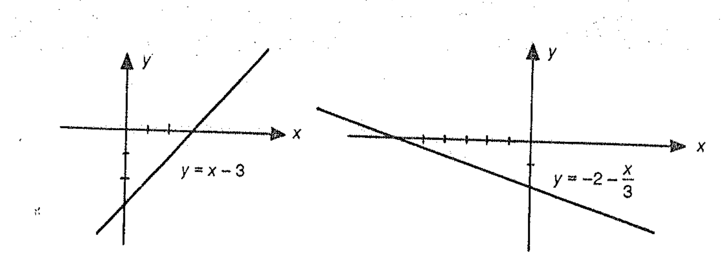
e)



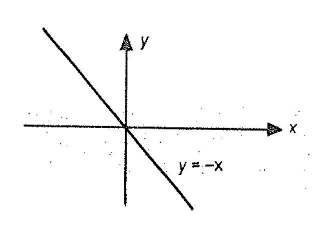
f)



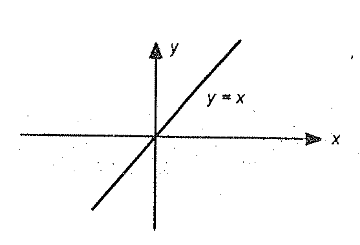
6) a) b)



c)



d), e), f)



10) a) $0 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3$
   b) $-1 - x - x^2$
   c) $3x^5 + x + 3$
   d) $-3x^3 - 24x^2 - 9x - 3$
   e) $x^2 + 2x + 5$
   f) $13x + 34$

11) a) $[0\ \ 0\ \ 0\ \ 0]$

b) $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

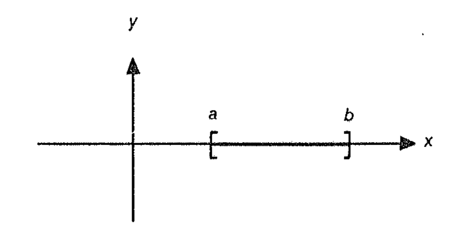
c) $\begin{bmatrix} 1 & -1 \\ -2 & -3 \end{bmatrix}$

d) $\begin{bmatrix} -2 & 6 & 14 \\ 12 & 2 & 2 \end{bmatrix}$

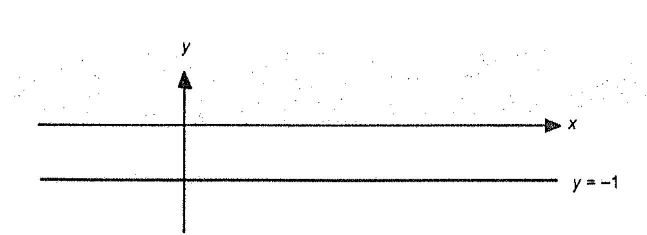
e) $[17\ \ -5]$

f) $\begin{bmatrix} -3 & 6 \\ -9 & -24 \\ -15 & 12 \end{bmatrix}$

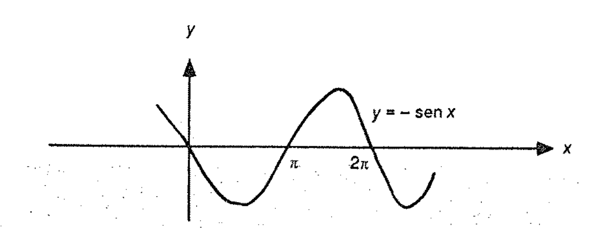
12) a)



b)



c)



d) Se debe cumplir que $c_1e^x + c_2xe^x = 0$ para toda x real. Al poner $x = 0$, se obtiene $c_1 = 0$. Al poner $x = 1$, se obtiene $c_2e = 0$, de donde $c_2 = 0$.

e) Al poner $c_1 = -14t$, $c_2 = 5t$ y $c_3 = t$ con $t \in \mathbf{R}$ $(t \neq 0)$ se obtiene $c_1(x + 1) + c_2(3x + 2) + c_3(-x + 4) = 0$.

SECCIÓN 3 (PÁGINAS 214-218)

2) a) Sí es subespacio
b) Sí es subespacio
c) Sí es subespacio
d) Sí es subespacio
e) No es subespacio
f) No es subespacio
g) No es subespacio
h) No es subespacio
i) Sí es subespacio

3) a) Sí es subespacio
b) Sí es subespacio
c) No es subespacio
d) Sí es subespacio
e) Sí es subespacio
f) Sí es subespacio
g) Sí es subespacio

4) a) Sí es subespacio
b) Sí es subespacio
c) Sí es subespacio
d) Sí es subespacio
e) Sí es subespacio
f) Sí es subespacio
g) Sí es subespacio
h) Sí es subespacio
i) Sí es subespacio
j) No es subespacio

7) a) $\left\{ (x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_1 = \frac{1}{2}t,\ x_2 = -\frac{1}{2}t,\ x_3 = t,\ t \in \mathbf{R} \right\}$

b) $\left\{ (x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_1 = x_2 = x_3 = t,\ t \in \mathbf{R} \right\}$

c) $\left\{ (x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_1 = s - r,\ x_2 = r,\ x_3 = s,\ r,\ s \in \mathbf{R} \right\}$

d) $\left\{ (x_1, x_2, x_3, x_4) \in \mathbf{R}^4 \mid x_1 = -t,\ x_2 = 6t,\ x_3 = 10t,\ x_4 = 3t,\ t \in \mathbf{R} \right\}$

8) Es el espacio trivial $\{0\}$.

9) No

10) $(2, 5) = \frac{12}{7}(1, 3) + \frac{1}{7}(2, -1)$

11) $\begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix} = -\frac{2}{11}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{9}{11}\begin{bmatrix} -1 & 3 \\ 0 & 4 \end{bmatrix} + \frac{8}{11}\begin{bmatrix} 0 & 1 \\ 3 & -1 \end{bmatrix} - \frac{37}{11}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

12) a) Sí es combinación lineal de $p_1(x)$ y $p_2(x)$: $-2 - x + x^2 = (x + x^2) - 2(1 + x)$

b) Sí es combinación lineal de $p_1(x)$ y $p_2(x)$: $3 + 2x - x^2 = -(x + x^2) + 3(1 + x)$

c) No es combinación lineal de $p_1(x)$ y $p_2(x)$.

d) No es combinación lineal de $p_1(x)$ y $p_2(x)$.

13) Todo $\mathbf{R}^2$

14) $\mathcal{L}(v_1, v_2) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2} \,\middle|\, a = \frac{1}{4}c - d,\ b = \frac{1}{8}c + 2d \right\}$

18) El subespacio trivial $\{0\}$.

19) a) $\{(x, y, z) \in \mathbf{R}^3 \mid x = t, y = 2t, z = -t, t \in \mathbf{R}\}$
b) $\{(x, y, z) \in \mathbf{R}^3 \mid z = 0\}$
c) $\{(x, y, z) \in \mathbf{R}^3 \mid 4x + 12y - 5z = 0\}$
d) $\{(x, y, z) \in \mathbf{R}^3 \mid x - 3y - z = 0\}$
e) Todo $\mathbf{R}^3$
f) $\{(x, y, z) \in \mathbf{R}^3 \mid 3x - 7y + 2z = 0\}$
g) Todo $\mathbf{R}^3$
h) $\{(x, y, z) \in \mathbf{R}^3 \mid x = t, y = t, z = t, t \in \mathbf{R}\}$

20) a) Sí pertenece a S d) No pertenece a S
b) No pertenece a S e) Sí pertenece a S
c) No pertenece a S f) Sí pertenece a S

21) Se tiene $1 = (1)\,\text{sen}^2 x + (1)\cos^2 x$, $\cos^2 x = (1)\cos^2 x + (-1)\,\text{sen}^2 x$. Entonces $1, \cos^2 x \in S$.

24) Dado un polinomio arbitrario $a_0 + a_1 x + \ldots + a_n x^n \in P_n$, éste es una combinación lineal de los vectores $1, x, \ldots, x^n$: el coeficiente del vector 1 es a_0, el coeficiente del vector x es a_1, etc.

25) Por ejemplo, el vector $x^n \in P_n$ no se encuentra en el espacio generado por los vectores $1, (x + 1), \ldots, (x + 1)^{n-1}$.

26) Por ejemplo, el vector $(0, 0, \ldots, 0, 1) \in \mathbf{R}^n$ no se encuentra en el espacio generado por los vectores $(1, 1, \ldots, 1, 0), (1, 1, \ldots, 1, 0, 0), \ldots, (1, 0, 0, \ldots, 0, 0)$.

27) Hay una infinidad de parejas de vectores de S que lo generan. Por ejemplo, $v_1 = (1, -1, -1)$ y $v_2 = (10, 20, 8)$, $v_1 = (20, 15, 1)$ y $v_2 (-10, 0, 4)$, etc. S no puede ser generado por un solo vector.

28) a) Falso: el vector $(0, 0, 0) \in \mathbf{R}^3$ no genera una recta.
b) Falso: si uno de los vectores es el vector $(0, 0, 0) \in \mathbf{R}^3$, éstos no generan un plano en $\mathbf{R}^3$.
c) Falso: por ejemplo, los vectores $v_1 = (1, 1, 1)$ y $v_2 = (2, 2, 2)$ son dos vectores no nulos de $\mathbf{R}^3$ y ellos generan la recta $x = t, t = t, z = t, t \in \mathbf{R}$.
d) Cierto.
e) Falso: vea los incisos d) y h) del ejercicio 19.
f) Cierto.

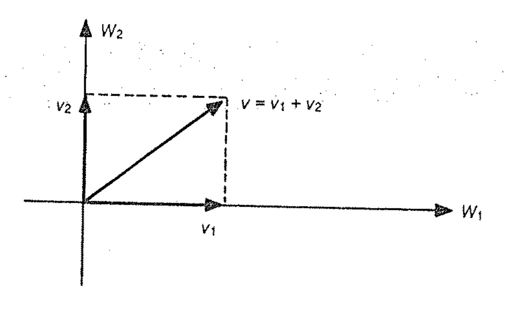
30) Se tienen los siguientes cinco casos:
I Si alguno de los subespacios es el subespacio trivial $\{0\}$, la intersección es el subespacio trivial $\{0\}$.
II Si alguno de los subespacios es $\mathbf{R}^3$, la intersección de éste con el otro subespacio coincide con este último.
III Si uno de los subespacios es la recta L que pasa por el origen y el otro un plano que pasa por el origen, la intersección es:
a) la recta L, si ésta se encontraba sobre el plano que representa el segundo subespacio, o
b) el subespacio $\{0\}$ si la recta L no se encontraba sobre el plano que representa dicho subespacio.
IV Si uno de los subespacios es la recta L_1 y el otro la recta L_2 (ambas pasando por el origen), su intersección es: a) $L_1 = L_2$ si las rectas coinciden, o b) $\{0\}$ si no coinciden.
V Si uno de los subespacios es el plano $\prod_1$ y el otro el plano $\prod_2$ (ambos pasando por el origen), su intersección es: a) $\prod_1 = \prod_2$ si los planos coinciden; o b) una recta que pasa por el origen, si los planos no coinciden.

31) Se tienen los siguientes cinco casos:
I Si alguno de los subespacios es $\{0\}$, la suma de éste con el otro subespacio coincide con este último.
II Si alguno de los subespacios es $\mathbf{R}^3$, la suma de éste con el otro subespacio es $\mathbf{R}^3$.
III Si uno de los subespacios es la recta L que pasa por el origen y el otro es el plano $\prod$ que pasa por el origen, entonces la suma es: a) el plano $\prod$ si la recta L se encontraba sobre éste; o b) $\mathbf{R}^3$ caso contrario.
IV Si uno de los subespacios es la recta L_1 y el otro la recta L_2 (ambas pasando por el origen) entonces su suma es: a) $L_1 = L_2$ si las rectas coinciden; o b) el plano en el que se encuentran L_1 y L_2 si las rectas no coinciden.
V Si uno de los subespacios es el plano $\prod_1$ y el otro el plano $\prod_2$ (ambos pasando por el origen) entonces, su suma es: a) $\prod_1 = \prod_2$ si los planos coinciden; o b) $\mathbf{R}^3$ si los planos no coinciden.

32) a) La recta $x = -\frac{3}{2}t$, $y = -\frac{1}{3}t$, $z = t \in \mathbf{R}$.
b) Por ejemplo, el vector $v_1 = (2, -1, -1) \in S_1 U S_2$ (de hecho $v_2 \in S_2$) son tales que $v_1 + v_2 = (2, 0, 0) \notin S_1 U S_2$.
d) Cualquier vector $(x, y, z) \in \mathbf{R}^3$ se puede escribir como la suma de los vectores $v_1 = (2x + y + z, y, -2x - 2y - z) \in S_1$

$$v_2 = (-x - y - z, 0, 2x + 2y + 2z) \in S_2$$

34) Todo vector $(x, y) \in \mathbf{R}^2$ se puede escribir como la suma del vector $v_1 = x(1, 0) \in W_1$ y el vector $v_2 = y(0, 1) \in W_2$. De aquí que $\mathbf{R}^2 = W_1 + W_2 = \mathcal{L}(W_1 U W_2)$. Geométricamente se tiene



SECCIÓN 4 (PÁGINAS 226-228)

2) Es una consecuencia directa del teorema 4.3: $v_1 = (a, b)$ y $v_2 = (c, d)$ son linealmente independientes si, y sólo si: $det\begin{bmatrix} a & c \\ b & d \end{bmatrix} = ad - bc \neq 0$.

3) a) es linealmente independiente
b) es linealmente independiente
c) es linealmente dependiente
d) es linealmente independiente
e) es linealmente dependiente
f) es linealmente dependiente
g) es linealmente dependiente
h) es linealmente dependiente
i) es linealmente independiente
j) es linealmente dependiente

4) Falso: si el vector es el vector cero, el conjunto es linealmente dependiente. La afirmación sería verdadera si dijera: cualquier conjunto formado por un solo vector *no nulo* es linealmente independiente.

5) Sí, y sólo si: $a_{11}a_{22} \ldots a_{nn} \neq 0$

6) $a = -2. (1, 2, 0, 0) = \frac{2}{5}(-1, 3, -2, 1) + \frac{4}{5}(4, 1, 1, 1) - \frac{3}{5}(3, 0, 0, 2)$

8) Los vectores v_1 y v_2 en $\mathbf{R}^2$ o $\mathbf{R}^3$ son linealmente independientes si, y sólo si no son colineales, o bien, v_1 no se encuentra sobre la recta (que pasa por el origen) en la que se encuentra v_2.

9) La ecuación del plano es $5x - 8y - 7z = 0$.

12) La expresión $c_1 + c_2(x + 2) + c_3(x + 2)^2 + c_4(x + 2)^3 = 0 \ (\in P_3)$ conduce al sistema

$$\begin{aligned} c_1 + 2c_2 + 4c_3 + 8c_4 &= 0 \\ c_2 + 4c_3 + 12c_4 &= 0 \\ c_3 + 6c_4 &= 0 \\ c_4 &= 0 \end{aligned}$$

cuya única solución es $c_1 = c_2 = c_3 = c_4 = 0$, como se quería.

14) Considere los cuatro polinomios $p_r(x) = \sum_{k=0}^{2} a_{rk}x^k \in P_2$

La expresión $\sum_{r=1}^{3} c_r p_r(x) = 0 \ (\in P_2)$

Conduce al sistema $\sum_{4=1}^{4} r_{rk}c_r = 0$ de 3 ecuaciones ($k = 0, 1, 2$) con 4 incógnitas (c_1, c_2, c_3, c_4) el cual se sabe que tiene soluciones no triviales (teorema 2.1, capítulo 1).

16) La combinación lineal $c_1u_1 + c_2u_2 + \ldots + c_nu_n = 0$ se escribe como $c_1v_1 + c_2(v_1 + v_2) + \ldots + c_n(v_1 + v_2 + \ldots + v_n) = 0$ o bien $(c_1 + c_2 + \ldots + c_n)v_1 + (c_2 + \ldots + c_n)v_2 + \ldots + c_nv_n = 0$. Por ser $v_1, v_2, \ldots, v_n$ vectores linealmente independientes, la última expresión implica que $c_1 + c_2 + \ldots + c_n = 0$, $c_2 + \ldots + c_n = 0, \ldots, c_n = 0$, de donde $c_1 = c_2 = \ldots = c_n = 0$, lo que muestra la independencia lineal de $u_1, u_2, \ldots, u_n$.

17) b) En este caso el wronskiano de $f_1(x), f_2(x), f_3(x)$ y $f_4(x)$ es un determinante de una matriz con la última línea de ceros (pues $f_i'''(x) = 0$, $i = 1, 2, 3, 4$). Por lo tanto $W(f_1, f_2, f_3, f_4)(x) = 0$.

c.1) Si $x \geq 0$, se tiene $W(f_1, f_2)(x) = det\begin{bmatrix} x^2 & 0 \\ 2x & 0 \end{bmatrix} = 0$

Si $x < 0$, se tiene $W(f_1, f_2)(x) = det\begin{bmatrix} 0 & x^2 \\ 0 & 2x \end{bmatrix} = 0$

Entonces $W(f_1, f_2) = 0 \quad \forall x \in \mathbf{R}$.

c.2) Se tiene que mostrar que la expresión $c_1f_1(x) + c_2f_2(x) = 0 \ \forall x \in \mathbf{R}$ implica que $c_1 = c_2 = 0$. Si $x \geq 0$, la expresión anterior obliga a que $c_1 = 0$ y si $x < 0$, se obtiene $c_2 = 0$. Entonces, $f_1(x)$ y $f_2(x)$ son linealmente independientes.

d) Para mostrar la independencia lineal de las funciones $f_1, f_2, \ldots, f_n$ basta mostrar que su wronskiano no es idénticamente nulo.

d.1) $W(e^{ax}, e^{bx})(x) = (b - a)w^{(a + b)x} \neq 0$

d.2) $W(e^x, xe^x, x^2e^x)(x) = 4e^{3x} \neq 0$

d.3) $W(\text{sen } x, \cos x)(x) = -1 \neq 0$

d.4) $W(e^x \text{ sen } x, e^x \cos x)(x) = -e^{2x} \neq 0$

SECCIÓN 5 (PÁGINAS 245-248)

1) a) Sí constituyen una base.
b) Sí constituyen una base.
c) Sí constituyen una base.
d) No constituyen una base.

2) a) Por ejemplo, una base de S es $\{(1, 0, -1), (0, 1, -1)\}$
b) Por ejemplo, una base de W es $\{(1, 0, 0, -1), (0, 1, 0, -1), (0, 0, 1, -1)\}$
(La respuesta no es única.)

3) a) $\beta = \{(1, -1)\}$
b) $\beta = \{(0, 1, 1)\}$
c) $\beta = \{(-7, -6, 1)\}$
d) $\beta = \{(-3, 2, 5, 3)\}$
e) $\beta = \{(2, 0, 4, 1)\}$
(La respuesta no es única.)

5) El conjunto vacío.

6) a) Dos vectores no pueden generar a $\mathbf{R}^3$.
b) Los vectores son linealmente dependientes.
c) Los vectores son linealmente dependientes.
d) Los vectores son linealmente dependientes.
e) Los vectores son linealmente dependientes.

7) a) Por ejemplo, $\beta = \{(1, 1, 1), (1, 0, 0), (0, 1, 0)\}$
b) Por ejemplo, $\beta = \{(0, 1, 0), (1, 0, 1), (1, 0, 2)\}$
c) Por ejemplo, $\beta = \{(2, 3, 1), (3, 1, 2), (1, 0, 0)\}$
(La respuesta en estos tres incisos no es única.)
d) $\beta = \{(1, 1, 1), (3, -1, 4), (0, 1, 2)\}$

8) a) No. Los vectores v_1 y v_2 son linealmente dependientes; cualquier conjunto que los contenga es linealmente dependiente.
b) No. Una base de $\mathbf{R}^3$ no puede tener más de 3 vectores.

9) b) Por ejemplo, $\beta = \{(2, 1, 1), (10, 0, 2), (0, 1, -1)\}$
(La respuesta no es única.)

10) a) Por ejemplo, $\beta = \{(0, 1, -3, 2), (1, -1, 0, 1), (3, 0, 1, -1)\}$
b) Por ejemplo, $\beta' = \{(0, 1, -3, 2), (1, -1, 0, 1), (3, 0, 1, -1), (0, 0, 0, 1)\}$
(La respuesta no es única.)

11) a) Se tiene que $v_3 = v_1 + 4v_2$ y $v_4 = 3v_1 + v_2$. Entonces, $\mathcal{L}(v_1, v_2, v_3, v_4) = \mathcal{L}(v_1, v_2)$, por lo que una base de W es $\beta = \{v_1, v_2\} = \{(1, -1, 1, 0, 3), (0, 3, -2, -1, 4)\}$
b) Por ejemplo, se obtiene $\beta' = \{v_1, v_2, (1, 0, 0, 0, 0), (0, 1, 0, 0, 0), (0, 0, 1, 0, 0)\}$
(La respuesta no es única.)

12) a) $\beta = \{v_1, v_3, v_4, v_5\}$ b) $\beta' = \{v_1, v_3, v_4, v_5, u_1, u_2\}$
con $u_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, con $u_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$
(La respuesta no es única.)

13) a) $\beta = \{1, x + 1, (x + 1)^2, \ldots, (x + 1)^n\}$
b) $\beta' = \beta$.

14) a) $W_1 + W_2 = M_{2\times 2}$, $W_1 \cap W_2 = \left\{\begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix} \,\middle|\, \alpha, \beta \in \mathbf{R}\right\}$
b) Una base de W_1 es $\beta_1 = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}$
Una base de W_2 es $\beta_1 = \left\{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$
Una base de $M_{2\times 2}$ es, por ejemplo, la base canónica de este espacio.
Una base de $W_1 \cap W_2$ es: $\beta_3 = \left\{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}$

c) Se tiene: $dim(W_1 + W_2) = dim\, M_{2\times 2} = 4$, $dim\, W_1 = 3$, $dim\, W_2 = 3$, $dim\, W_1 \cap W_2 = 2$.
Entonces, $dim(W_1 + W_2) = dim\, W_1 + dim\, W_2 - dim\,(W_1 \cap W_2)$

15) a) $W_1 + W_2 = \left\{\begin{bmatrix} \alpha & 0 & \beta \\ 0 & \alpha+\beta & \gamma \\ \beta & \gamma & \delta \end{bmatrix} \,\middle|\, \alpha, \beta, \gamma, \delta \in \mathbf{R}\right\}$

$W_1 \cap W_2 = \left\{\begin{bmatrix} 0 & 0 & \alpha \\ 0 & \alpha & 0 \\ \alpha & 0 & 0 \end{bmatrix} \,\middle|\, \alpha \in \mathbf{R}\right\}$

b) Una base de W_1 es: $\beta_1 = \left\{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right\}$

Una base de W_2 es $\beta_2 = \left\{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right\}$

Una base de $W_1 + W_2$ es

$$\beta_3 = \left\{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right\}$$

Una base de $W_1 \cap W_2$ es $\beta_4 = \left\{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\right\}$

c) Se tiene: $dim\,(W_1 + W_2) = 4$, $dim\, W_1 = 2$, $dim\, W_2 = 3$, $dim\, W_1 \cap W_2 = 1$. Entonces $dim\,(W_1 + W_2) = dim\, W_1 + dim\, W_2 - dim\,(W_1 \cap W_2)$.

16) a) Una base para W_1 está formada por las $\frac{n(n+1)}{2}$ matrices

$$\beta_1 = \left\{\begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \\ & & \ldots & & \\ 1 & 0 & 0 & \ldots & 0 \end{bmatrix}, \right.$$

$$\begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 1 \\ & & \ldots & & \\ 0 & 1 & 0 & \ldots & 0 \end{bmatrix}$$

$$\left. \ldots, \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}\right\}$$

Una base para W_2 está dada por las $\frac{n(n-1)}{2}$ matrices

$$\beta_2 = \left\{ \begin{bmatrix} 0 & 1 & 0 \dots 0 \\ -1 & 0 & 0 \dots 0 \\ & \cdot \; \cdot \; \cdot & \\ 0 & 0 & 0 \dots 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \dots 0 \\ 0 & 0 & 0 \dots 0 \\ -1 & 0 & 0 \dots 0 \\ & \cdot \; \cdot \; \cdot & \\ 0 & 0 & 0 \dots 0 \end{bmatrix}, \dots, \right.$$

$$\begin{bmatrix} 0 & 0 & 0 \dots 1 \\ 0 & 0 & 0 \dots 0 \\ & \cdot \; \cdot \; \cdot & \\ -1 & 0 & 0 \dots 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \dots 0 \\ 0 & 0 & 1 \dots 0 \\ 0 & -1 & 0 \dots 0 \\ & \cdot \; \cdot \; \cdot & \\ 0 & 0 & 0 \dots 0 \end{bmatrix}, \dots,$$

$$\left. \begin{bmatrix} 0 & 0 & 0 \dots 0 \\ 0 & 0 & 0 \dots 1 \\ & \cdot \; \cdot \; \cdot & \\ 0 & -1 & 0 \dots 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & 0 & 0 \dots & 0 & 0 \\ 0 & 0 & 0 \dots & 0 & 0 \\ & & \cdot \; \cdot \; \cdot & & \\ 0 & 0 & 0 \dots & 0 & 1 \\ 0 & 0 & 0 \dots & -1 & 0 \end{bmatrix} \right\}$$

c) Sea A una matriz de orden n y k un escalar. Entonces $(k(A + A^t))^t = k(A^t + A)$ y $(k(A - A^t))^t = k(A^t - A) = -k(A - A^t)$ por lo que $k(A + A^t)$ es una matriz simétrica, mientras que $k(A - A^t)$ es una matriz antisimétrica. Observe que $A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t)$ lo que muestra que $M_{n\times n} = W_1 + W_2$. Suponga ahora que $A \in W_1 \cap W_2$. Entonces, $A = A^t = -A$, de donde $A = 0$, esto es, $W_1 \cap W_2 = \{0\}$. Entonces, $M_{n\times n} = W_1 \oplus W_2$.

d) Se tiene $n^2 = \dfrac{n(n+1)}{2} + \dfrac{n(n-1)}{2}$

$$\begin{matrix} \| & \| & \| \\ dim\, M_{n\times n} & dim\, W_1 & dim\, W_2 \end{matrix}$$

17) Sean v_1 y v_2 dos vectores no colineales en el plano P_1 y v_3 cualquier vector no nulo en el plano P_2. Es claro que $\beta_1 = \{v_1, v_2\}$ es una base de P_1. También debe ser claro que $\beta = \{v_1, v_2, v_3\}$ es una base de $\mathbf{R}^3$, de modo que todo vector $(x, y, z) \in \mathbf{R}^3$ se puede escribir como la suma de un vector $u_1 = c_1v_1 + c_2v_2 \in P_1$ y un vector $u_2 = c_3v_3 \in P_2$. Esto muestra que $\mathbf{R}^3 = P_1 + P_2$. Por otra parte, del teorema 5.6 se concluye que $dim(P_1 \cap P_2) = dim\, P_1 + dim\, P_2 - dim(P_1 + P_2) = 2 + 2 - 3 = 1$, por lo que $P_1 \cap P_2$ es una recta que pasa por el origen.

18) Falso. Considere por ejemplo, $V = M_{2\times 3}$

$$v_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\quad v_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},\quad v_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$w_1 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},\quad w_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$u_1 = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix},\quad u_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

En este caso, $W_1 = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \in M_{2\times 3} \mid f = 0 \right\}$

$$W_2 = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \in M_{2\times 3} \mid e = 0 \right\}$$

$$W_1 \cap W_2 = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \in M_{2\times 3} \mid e = f = 0 \right\}$$

de modo que una base de $W_1 \cap W_2$ es $\{v_1, v_2, v_3, w_1\}$ (o $\{v_1, v_2, v_3, u_1\}$) y no $\{v_1, v_2, v_3\}$.

SECCIÓN 6 (PÁGINAS 264-266)

1) a) Si $x = (x_1, x_2, \dots, x_n) \in \mathbf{R}^n$, entonces, $(x)_\beta = (x_1, x_2, \dots, x_n)$.

b) Si $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ entonces,

$(A)_\beta = (a_{11}, a_{12}, \dots, a_{1n}, a_{21}, a_{22}, \dots, a_{2n}, \dots, a_{m1}, a_{m2}, \dots, a_{mn}) \in \mathbf{R}^{mn}$

c) Si $P(x) = a_0 + a_1x + \dots + a_nx^n \in P_n$, entonces $(P)_\beta = (a_0, a_1, \dots, a_n) \in \mathbf{R}^{n+1}$

2) a) El vector 0 de $\mathbf{R}^n$.

b) El i-ésimo vector e_i de la base canónica de $\mathbf{R}^n$, $i = 1, 2, \dots, n$.

3) a) β es una base de $\mathbf{R}^3$ puesto que $det \begin{bmatrix} 3 & -1 & 4 \\ 1 & 0 & 3 \\ 2 & 2 & 5 \end{bmatrix} = -11 \neq 0$

(véase teorema 4.3).

b) $((1, 1, 1))_\beta = \left(-\frac{1}{4}, -\frac{3}{11}, \frac{5}{11}\right)$ $((3, 4, 6))_\beta = \left(-\frac{16}{11}, \frac{-1}{11}, \frac{20}{11}\right)$

c) $((x, y, z))_\beta = (c_1, c_2, c_3)$ en donde: $c_1 = \frac{1}{11}(6x - 13y + 3z)$,

$c_2 = \frac{1}{11}(-x - 7y + 5x)$, $c_3 = \frac{1}{11}(-2x + 8y - z)$

4) b) $(1)_\beta = \left(\frac{3}{7}, -\frac{1}{14}, -\frac{1}{2}, \frac{5}{14}\right)$, $(x)_\beta = \left(-\frac{2}{7}, \frac{3}{14}, \frac{1}{2}, -\frac{1}{14}\right)$

$(x^2)_\beta = \left(0, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right)$, $(x^3)_\beta = \left(\frac{6}{7}, -\frac{9}{14}, -\frac{1}{2}, \frac{3}{14}\right)$

6) $(a_0 + a_1x + a_2x^2 + a_3x^3)_\beta = (c_1, c_2, c_3, c_4)$ en donde:

$c_1 = \frac{3}{7}a_0 - \frac{2}{7}a_1 + \frac{6}{7}a_3$, $c_2 = -\frac{1}{14}a_0 + \frac{3}{14}a_1 + \frac{1}{2}a_2 + \frac{3}{14}a_3$

$c_3 = -\frac{1}{2}a_0 + \frac{1}{2}a_1 + \frac{1}{2}a_2 - \frac{1}{2}a_3$, $c_4 = \frac{5}{14}a_0 - \frac{1}{14}a_1 - \frac{1}{2}a_2 + \frac{3}{14}a_3$

7) $v_1 = (8, 2)$, $v_2 = (-5, -1)$

8) $v_1 = 1 - \frac{2}{5}x + \frac{3}{5}x^2$, $v_2 = -1 + \frac{9}{5}x - \frac{6}{5}x^2$, $v_3 = 2 - \frac{12}{5}x + \frac{8}{5}x^2$

9) $(p)_\beta = (c_0, c_1, \dots, c_n) \in \mathbf{R}^{n+1}$, en donde $c_i = \frac{1}{i!}p^{(i)}(1)$, $i = 0, 1, \dots, n$

10) a) $((1, 1, 3))_{\beta_1} = (3, -2, 0)$

b) $((1, 1, 3))_{\beta_2} = (1/17, 19/17, 4/17)$

c) $P = \begin{bmatrix} 9/17 & 13/17 & 12/17 \\ 1/17 & -8/17 & -10/17 \\ 2/17 & 1/17 & -3/17 \end{bmatrix}$

d) $P^{-1} = \begin{bmatrix} 2 & 3 & -2 \\ -1 & -3 & 6 \\ 1 & 1 & -5 \end{bmatrix}$

e) $\begin{bmatrix} 2 & 3 & -2 \\ -1 & -3 & 6 \\ 1 & 1 & -5 \end{bmatrix} \begin{bmatrix} 1/17 \\ 19/17 \\ 4/17 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 0 \end{bmatrix}$

$\begin{bmatrix} 9/17 & 13/17 & 12/17 \\ 1/17 & -8/17 & -10/17 \\ 2/17 & 1/17 & -3/17 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/17 \\ 19/17 \\ 4/17 \end{bmatrix}$

11) a) $P = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 4 \end{bmatrix}$

b) $P^{-1} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1/2 & -1/2 & 0 \\ 0 & 0 & 1/3 & -1/3 \\ 0 & 0 & 0 & 1/4 \end{bmatrix}$

c) $((a, b, c, d))_{\beta_2} = \left(a - b, \frac{1}{2}(b - c), \frac{1}{3}(c - d), \frac{1}{4} d\right)$

12) a) $P = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$

b) $P^{-1} = P$

c) $((x_1, x_2, x_3, x_4, x_5))_{\beta_2} = (x_5, x_4, x_3, x_2, x_1)$

13) $P = \begin{bmatrix} 3 & 3 & -4 & -3 \\ 0 & 6 & 1 & 1 \\ 5 & 4 & 2 & 1 \\ 2 & 3 & 3 & 2 \end{bmatrix}$

14) c) $P = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$ d) $P^{-1} = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$

15) c) $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ d) $P = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$

16) a) Se intercambian de posición dos de sus columnas (las correspondientes a los vectores intercambiados de β_1).

b) Se intercambian de posición dos de sus líneas (las correspondientes a los vectores intercambiados de β_2).

17) $f\colon P_1 \to \mathbf{R}^3$, $f(a_0 + a_1x + a_2x^2)$

$= \left(\frac{2}{3} a_0 - \frac{1}{4} a_1 - \frac{7}{12} a_2,\ \frac{1}{3} a_0 + \frac{1}{3} a_2,\ \frac{1}{3} a_0 - \frac{1}{4} a_1 + \frac{1}{12} a_2\right)$

18) $f\colon M_{2\times 2} \to R^4, f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \frac{1}{4}\ (a + b + c + d, a + b - c - d,\ a - b + c - d,$
$a - b - c + d)$

$f^{-1}(x_1, x_2, x_3, x_4) = \frac{1}{4}\begin{bmatrix} x_1 + x_2 + x_3 + x_4 & x_1 + x_2 - x_3 - x_4 \\ x_1 - x_2 + x_3 - x_4 & x_1 - x_2 - x_3 + x_4 \end{bmatrix}$

19) $f\colon P_3 \to M_{2\times 2}, f(a_0 + a_1x + a_2x^2 + a_3x^3) = \begin{bmatrix} a_0 & a_1 \\ a_2 & a_3 \end{bmatrix}$

SECCIÓN 7 (PÁGINAS 275-276)

2) Todo $\mathbf{R}^n$.

3) Por ejemplo, $\beta = \{1, 0, 4/5, -2/5), (0, 1, -3/5, 4/5)\}$

4) a) dimensión = 2
b) dimensión = 1
c) dimensión = 1
d) dimensión = 2
e) dimensión = 3
f) dimensión = 4

5) a) $\beta = \left\{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\right\}$

b) $\beta = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 24/33 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 23/22 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & -2/11 \end{bmatrix}\right\}$

c) $\beta = \left\{\begin{bmatrix} 1 & 0 \\ 12/5 & 4/5 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -3/5 & -1/5 \end{bmatrix}\right\}$

d) $\beta = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 3 & 1/3 \end{bmatrix}\right\}$

(La respuesta no es única, la que se proporciona es en atención al teorema 7.2.)

6) a) $\beta = \left\{1 + \frac{4}{3}x^2, x + \frac{1}{3}x^2\right\}$

b) $\beta = \left\{1 - \frac{1}{11}x^2, x + \frac{13}{11}x^2\right\}$

c) $\beta = \left\{1 - \frac{4}{9}x^2, x + \frac{5}{9}x^2\right\}$

d) $\beta = \left\{1 - \frac{2}{11}x^2, x - \frac{1}{11}x^2\right\}$

(Véase aclaración en el problema anterior.)

8) Se tiene de hecho $W_1 \subset W_2$, de modo que $W_1 \cap W_2 = W_1$. Una base de $W_1 = W_1 \cap W_2$ es: $\beta = \{(8, 12, 1, 0), (10, 15, 0, 1)\}$. Una base de $W_1 + W_2$ es: $\beta = \{(1, 3/2, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$.

9) Una base de $W_1 \cap W_2$ es: $\beta = \{(1, 0, 1, 0), (0, 1, 0, 1)\}$. Una base de $W_1 + W_2$ es: $\beta = \{(1, 0, 0, 0), (0, 1, 0, 1), (0, 0, 1, 0)\}$.

10) a) $x_1 - 2x_2 = 0$, $x_3 - 3x_2 = 0$
b) $5x_1 + 2x_2 - 3x_3 = 0$
c) $4x_1 + 2x_2 - 3x_3 = 0$, $x_1 - 2x_2 - 3x_4 = 0$
d) $x_1 - 2x_3 + 3x_4 = 0$, $2x_1 - x_3 + 6x_5 = 0$

CAPÍTULO CUATRO

SECCIÓN 1 (PÁGINAS 278-285)

1) a) Sí es lineal. b) Sí es lineal. c) Sí es lineal. d) No es lineal. e) No es lineal. f) No es lineal.

2) a) Sí es lineal. b) Sí es lineal. c) No es lineal. d) Sí es lineal. e) Sí es lineal. f) Sí es lineal.

3) a) Sí es lineal. b) Sí es lineal. c) Sí es lineal. d) No es lineal. e) Sí es lineal. f) No es lineal.

4) a) Sí es lineal. b) No es lineal. c) Sí es lineal. d) Sí es lineal. e) Sí es lineal. f) Sí es lineal.

5) T es lineal puesto que $T(A_1 + A_2) = (A_1 + A_2)^t = A_1^t + A_2^t = T(A_1) + T(A_2)$, $T(cA) = (cA)^t = cA^t = cT(A)$.
(Véase ejercicio 14 de la sección 2, capítulo 2.)

6) Si A conmuta con B, entonces $T(A) = 0$. Dos ejemplos triviales de matrices de este tipo son la matriz 0 y la matriz I.

7) b) $T: M_{n\times n} \to \mathbf{R}$, $T(A) = \det A$, no es una transformación lineal.

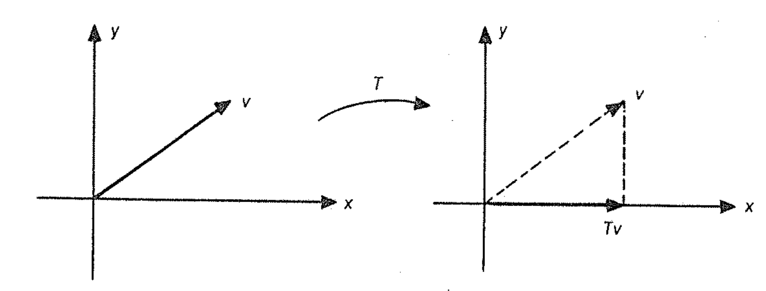
11) Sea $\beta = \{v_0\}$ ($v_0 \in V$, $v_0 \neq 0$) una base de V, considere la imagen bajo T del vector v_0, $T(v_0) \in V$. Existe un escalar $c \in \mathbf{R}$ tal que $T(V_0) = cv_0$. Para cualquier vector $v \in V$, existe $k \in \mathbf{R}$ tal que $v = kv_0$. Entonces, $T(v) = T(kv_0) = kT(v_0) = k(cv_0) = c(kv_0) = cv$.

12) Tome la combinación lineal $c_1v_1 + c_2v_2 + \ldots + c_kv_k = 0$. Entonces, ya que T es lineal $c_1T(v_1) + \ldots + c_nT(v_n) = 0$. Como los vectores $T(v_1), T(v_2), \ldots, T(v_k)$ son linealmente independientes, se concluye que $c_1 = c_2 = \ldots = c_k = 0$, lo que muestra la independencia lineal de los vectores $v_1, v_2, \ldots, v_k$. Este argumento *NO* es reversible (¿por qué?). La afirmación recíproca es en general falsa. Por ejemplo, considere la transformación lineal del inciso f), ejercicio 2: $T: \mathbf{R}^2 \to \mathbf{R}$, $T(x, y) = (0, y)$. Los vectores $v_1 = (1, 1)$ y $v_2 = (1, -1)$ son linealmente independientes, pero los vectores $T(v_1) = (0, 1)$ y $T(v_2) = (0, -1)$ son linealmente dependientes.

13) Dado $u \in V$, existen escalares $c_1, c_2, \ldots, c_k$ tales que $c_1T(v_1) + c_2T(v_2) + \ldots + c_kT(v_k) = u$, expresión que puede reescribirse como $T(c_1v_1 + c_2v_2 + \ldots + c_kv_k) = u$. Tome $v = c_1v_1 + c_2v_2 + \ldots + c_kv_k \in V$. Entonces $T(v) = u$, lo que muestra que T es sobreyectiva.

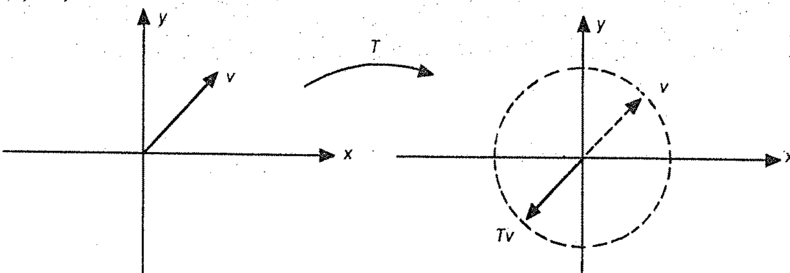
15) b) Si el vector $v \in \mathbf{R}^2$ se encuentra sobre el eje y, entonces $T(v) = v$.

16) b)



c) Si el vector $v \in \mathbf{R}^2$ se encuentra sobre el eje y, entonces $T(v) = 0$.
Si el vector $u \in \mathbf{R}^2$ se encuentra sobre el eje x, entonces $T(u) = u$.

17) b)



18) a) Suponga que existe otra transformación lineal $\tilde{T}: \mathbf{R}^2 \to \mathbf{R}^2$ tal que $\tilde{T}((1, 0)) = (2, 3)$ y $\tilde{T}((0, 1)) = (1, 1)$. Para cualquier vector $(x, y) \in \mathbf{R}^2$ se puede escribir $(x, y) = x(1, 0) + y(0, 1)$. Entonces, $\tilde{T}((x, y)) = \tilde{T}(x(1, 0) + y(0, 1)) = x\tilde{T}((1, 0)) + y\tilde{T}((0, 1)) = x(2, 3) + y(1, 1) = xT((1, 0)) + yT((0, 1)) = T(x(1, 0) + y(0, 1)) = T((x, y))$, lo que muestra que $\tilde{T} = T$.

b) $T((3, 4)) = T(3(1, 0) + 4(0, 1)) = 3T((1, 0)) + 4T((0, 1)) = 3(2, 3) + 4(1, 1) = (10, 13)$

c) $T((x, y)) = (2x + y, 3x + y)$

19) $T((3, 2)) = T(-4(2, 5) + 11(1, 2)) = -4T((2, 5)) + 11T((1, 2)) = -4(3, 1) + 11(0, 3) = (-12, 29)$

20) $T((2, 4, -2)) = T(28(1, 1, 1) - 13(2, 0, 0) - 6(0, 4, 5)) = 28(1 - x + x^2) - 13(3 + x - x^2) - 6(2 + 3x - x^2) = -23 - 59x + 47x^2$

21) Demostración del corolario del teorema 1.1: Dado $v \in V$ se puede escribir $v = c_1v_1 + c_2v_2 + \ldots + c_nv_n$. Entonces, $T_1(v) = T_1(c_1v_1 + c_2v_2 + \ldots + c_nv_n) = c_1T_1(v_1) +$

$c_2T_1(v_2) + \ldots + c_nT_1(v_n) = c_1T_2(v_1) + c_2T_2(v_2) + \ldots + c_nT_2(v_n) = T_2(c_1v_1 + c_2v_2 + \ldots + c_nv_n) = T_2(v)$, lo que muestra que $T_1 = T_2$.

22) Es una consecuencia inmediata del corolario del teorema 1.1, pues la transformación cero es tal que $0(v_j) = 0, j = 1, 2, \ldots, n$.

23) Es una consecuencia inmediata del corolario del teorema 1.1, pues el operador identidad en V es tal que $Id(v_j) = v_j,\ j = 1, 2, \ldots, n$.

SECCIÓN 2 (PÁGINAS 288-296)

1) Para la transformación lineal $T: V \to U$, al menos el cero de V se encuentra en *Ker T* y al menos el cero de U se encuentra en *ImT*.

2) *Ker T* = {polinomios constantes}

3) La función $f \in C([-\pi, \pi])$ está en el núcleo de T si, y sólo si $\int_{-\pi}^{\pi} f(x)dx = 0$. Por ejemplo, todas las funciones impares en $[-\pi, \pi]$ pertenecen a *Ker T*.

4) $Ker\ T = V,\ ImT = \{0\}$

5) $Ker\ T = \{0\},\ ImT = V.$

6) Se tiene $dim\ ImT \leq dim\ \mathbf{R} = 1$, esto es, $dim\ ImT = 0$ o 1. Como existe $v \in V, v \neq 0$ tal que $Tv \neq 0$, se concluye que $dim\ ImT = 1$. Por el teorema de la dimensión $dim\ Ker\ T = 0$. De aquí se sigue la conclusión del ejercicio.

7) $Ker\ T = \{(x, y) \mid y = x\}$ $\qquad ImT = \{(x, y) \mid y = x\}$

8) La transformación lineal es $T((x, y)) = \frac{1}{7}(3y - 2x)v$, en donde $v = T((1, 3))$. Entonces, como $v \neq 0, (x, y) \in Ker\ T \Leftrightarrow \frac{1}{7}(3y - 2x)\ v = 0 \Leftrightarrow 3y = 2x$, esto es, el núcleo de T es la recta $3y = 2x$.

9) La transformación lineal es $T((x, y, z)) = \frac{1}{5}(x - y + 2z)v$, en donde $v = T((-1, 2, 4)) \neq 0$. Entonces, $Ker\ T = \{(x, y, z) \mid x - y + 2z = 0\}$.

10) *Ker T* = espacio solución de $AX = 0$.

11) $Ker\ T = \left\{\begin{bmatrix} \alpha & \beta \\ 0 & \alpha \end{bmatrix} \mid \alpha, \beta \in \mathbf{R}\right\}$.

14) Se tiene $dim\ Ker\ T = 1$. Entonces $dim\ ImT = dim\ \mathbf{R}^2 - dim\ Ker\ T = 2 - 1 = 1$, por lo que *ImT* es también una recta que pasa por el origen.

16) Al ser *ImT* un subespacio propio de V, se tiene $dim\ ImT < dim\ V$. Entonces, $dim\ Ker\ T = dim\ V - dim\ ImT > 0$. Se sigue entonces la conclusión del ejercicio.

18) a) $Ker\ T = \{0\},\ ImT = \mathbf{R}^2,\ dim\ Ker\ T + dim\ ImT = 0 + 2 = dim\ \mathbf{R}^2$

b) Base de $Ker\ T = \{(1, 0, 1)\}$. Base de $ImT = \{(1, 0), (0, 1)\}$
$dim\ Ker\ T + dim\ ImT = 1 + 2 = 3 = dim\ \mathbf{R}^3$.

c) Base de $Ker\ T = \{(-1, 0, 1), (-1, 1, 0)\}$. Base de $ImT = \{(1, 2)\}$
$dim\ Ker\ T + dim\ ImT = 2 + 1 = 3 = dim\ \mathbf{R}^3$

d) $Ker\ T = \{0\}$. Base de $ImT = \{(1, 5, 0), (0, 0, 1)\}$
$dim\ Ker\ T + dim\ ImT = 0 + 2 = 2 = dim\ \mathbf{R}^2$.

e) $Ker\ T = \{0\},\ ImT = \mathbf{R}^3$
$dim\ Ker\ T + dim\ ImT = 0 + 3 = 3 = dim\ \mathbf{R}^3$.

f) Base de $Ker\ T = \{(2, -1, 1, 0), (3, -1, 0, 1)\}$
Base de $ImT = \{(1, 0, 1, 5), (0, 1, -3, -12)\}$
$dim\ Ker\ T + dim\ ImT = 2 + 2 = 4 = dim\ \mathbf{R}^4$.

g) $Ker\ T = \{0\}$. Base de $ImT = \{(1, 0, -1, -2), (0, 1, 2, 3)\}$
$dim\ Ker\ T + dim\ ImT = 0 + 2 = 2\ dim\ \mathbf{R}^2$.

h) Base de $Ker\ T = \{1 - x\}$. Base de $ImT = \{1, x, x^2\}$
$dim\ Ker\ T + dim\ ImT = 1 + 3 = 4 = dim\ P_3$

i) $Ker\ T = \{0\}$. Base de $ImT = \{x, x^2, x^3\}$
$dim\ Ker\ T + dim\ ImT = 0 + 3 = 3 = dim\ P_2$.

j) $Ker\ T = \{0\}$. Base de $ImT = \{x^2, x^3, x^4\}$
$dim\ Ker\ T + dim\ ImT = 0 + 3 = 3 = dim\ P_2$

k) $Ker\ T = \{0\},\ ImT = M_{2\times 2}$
$dim\ Ker\ T + dim\ ImT = 0 + 4 = 4 = dim\ M_{2\times 2}$

l) Base de $Ker\ T = \left\{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right\}$
$ImT = \mathbf{R}$
$dim\ Ker\ T + dim\ ImT = 8 + 1 = 9 = dim\ M_{3\times 3}$.

m) $Ker\ T = \{0\}$
Base de $ImT = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\right\}$
$dim\ Ker\ T + dim\ ImT = 0 + 4 = 4 = dim\ M_{2\times 2}$

n) $Ker\ T = \{0\}\quad ImT = P_3$
$dim\ Ker\ T + dim\ ImT = 0 + 4 = 4 = dim\ P_3$

SECCIÓN 3 (PÁGINAS 299-316)

1) a) Una base de *Ker T* es $\{(1, 0, 2), (0, 1, 0)\}$
Una base de *ImT* es $\{(1, -2)\}$

b) Una base de *Ker T* es $\{(1, -1, 1, 0), (-1, -4, 0, 2)\}$
Una base de *ImT* es $\{(1, 0, -3/2), (0, 1, 1/2)\}$

c) Una base de *Ker T* es $\{(1, 5, 2, -5)\}$
Una base de *ImT* es $\{(1, 0, -1, 0, 1), (0, 1, 1/2, 0, -1/2), (0, 0, 0, 1, -1/2)\}$

2) a) $\begin{bmatrix} 3 & -2 \\ 5 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 5 & 2 & -1 \\ 3 & 4 & 2 \end{bmatrix}$

c) $\begin{bmatrix} 2 & 1 & -1 \\ 3 & 2 & -2 \\ 8 & -1 & 0 \end{bmatrix}$

d) $\begin{bmatrix} 2 & 1 \\ 1 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

e) $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

f) $[\,1 \quad 1 \quad 2 \quad -3 \quad 8\,]$

g) $\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$

h) $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

i) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

j) $\begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 0 & -2 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

k) $\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$

l) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

m) $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

n) $[\,1 \quad 1 \quad 1 \quad 1\,]$

o) $[\,1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1\,]$

3) La matriz cero de orden $(\dim U) \times \dim V)$

5) a) $Ker\ T = \{0\}$, $Im T = \mathbf{R}^2$

b) Una base de $Ker\ T$ es $\{(0, 1, 2)\}$, $Im T = \mathbf{R}^2$

c) $Ker\ T = \{0\}$, $Im T = \mathbf{R}^3$

d) $Ker\ T = \{0\}$. Una base de $Im T$ es: $\{(1, 0, 3/7, 1/7), (0, 1, 1/7, -2/7)\}$

e) Una base de $Ker\ T$ es $(-1, 1, 0, 0), (0, 0, -1, 1)$, $Im T = \mathbf{R}^2$

f) Una base de $Ker\ T$ es $\{(-1, 1, 0, 0, 0), (-2, 0, 1, 0, 0), (3, 0, 0, 1, 0), (-8, 0, 0, 0, 1)\}$. $Im T = \mathbf{R}$

g) Una base de $Ker\ T$ es $\{(-1, 0, 0, 1, 0, 0), (0, -1, 0, 0, 1, 0), (0, 0, -1, 0, 0, 1)\}$ $Im T = \mathbf{R}^3$

h) $Ker\ T = \{0\}$. Una base de $Im T$ es: $\{x, x^2, x^3\}$

i) $Ker\ T = \{0\}$, $Im T = P_1$

j) Una base de $Ker\ T$ es $\{2 - 5x + x^2\}$, $Im T = \mathbf{R}^3$.

k) $Ker\ T = \{0\}$. Una base de $Im T$ es $\{1 - x^4, x + x^2 + x^3 + x^4\}$

l) $Ker\ T = \{0\}$. $Im T = M_{2\times 2}$.

m) $Ker\ T = \{0\}$. Una base de $Im T$ es:

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} \right\}$$

n) Una base de $Ker\ T$ es : $\left\{ \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

$Im T = \mathbf{R}$

o) Véase inciso 1) del ejercicio 18 en la sección anterior.

6) a) La matriz de la transformación T respecto de la base canónica de $M_{2\times 2}$ es:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Al describir el espacio solución del sistema $AX = 0$ (y luego regresando a $M_{2\times 2}$) se llega a la conclusión de que el núcleo de T es:

$$Ker\ T = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2} \mid b = c \right\}$$

que son precisamente las matrices simétricas de $M_{2\times 2}$.

Similarmente, el espacio línea de A^t queda descrito por las matrices

$$Im T = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2} \mid b = -c \right\}$$

que son precisamente las matrices antisimétricas de $M_{2\times 2}$.

7) $[T]_\beta = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 6 & 0 \\ 0 & 3 & 0 & 6 \end{bmatrix}$

Una base de $Ker\ T$ es: $\left\{ \begin{bmatrix} -2 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -2 \\ 0 & 1 \end{bmatrix} \right\}$

Una base de ImT es: $\left\{\begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 3 \end{bmatrix}\right\}$

8) La matriz C es: $C = \begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}$

Para calcular $det\ C$ se puede usar el resultado del ejercicio 13 de la sección 4 del capítulo 2:

$$det\ C = det \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = (ad - bc)^2$$

Se tiene, entonces, que la matriz B es inversible si, y sólo si la matriz C lo es. Para que exista $A \in M_{2\times 2}, A \neq 0$ tal que $T(A) = 0$ es necesario y suficiente que el sistema $CX = 0$ tenga soluciones no triviales (en donde $X = [A]_\beta$, β la base canónica de $M_{2\times 2}$). Esto ocurre si, y sólo si C no es inversible, esto es, si, y sólo si B es no inversible.

9) Véase solución del ejercicio 12 inciso c.3) de la sección 3 del capítulo 1.

10) c) $[T]_\beta = \begin{bmatrix} I_k & 0_{kxr} \\ 0_{rxk} & 0_r \end{bmatrix}$

d) $Ker\ T = W_2$, $ImT = W_1$

11) c) $[T] = \begin{bmatrix} I_{dim\ V_1} & & & 0 \\ & I_{dim\ V_2} & & \\ & & kI_{dim\ V_3} & \\ 0 & & & kI_{dim\ V_4} \end{bmatrix}$

12) a) $[D]_\beta = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

b) $[D]_\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

c) $[D]_\beta = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

d) $[D]_\beta = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

e) $[D]_\beta = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

f) $[D]_\beta = \begin{bmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

g) $[D]_\beta = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

13) a) Se intercambian las columnas i y j.

b) Se intercambian las líneas i y j.

14) a) $[T]_{\beta_1} = \begin{bmatrix} 2 & 0 & 1 & 1 \\ 3 & 1 & 2 & -1 \\ 2 & 0 & -1 & 4 \\ 2 & 3 & 4 & 1 \end{bmatrix}$

b) $[T]_{\beta_2} = \begin{bmatrix} 1 & 2 & 0 & 1 \\ -3 & -3 & -3 & -6 \\ 3 & 3 & 0 & 5 \\ 0 & 2 & 6 & 5 \end{bmatrix}$

c) $[T]_{\beta_3} = \begin{bmatrix} 4 & 2 & 1 & 2 \\ -2 & 1 & 1 & -1 \\ 7 & 2 & 4 & 5 \\ -5 & 4 & -1 & -6 \end{bmatrix}$

15) a) $[T]_{\beta_1\beta_2} = \begin{bmatrix} 2 & 1 & 1 \\ -3 & 1 & 0 \\ 1 & 1 & -2 \\ -1 & 1 & 0 \end{bmatrix}$

b) $[T]_{\beta_1'\beta_2'} = \begin{bmatrix} 3 & 20/3 & -1/6 \\ 1 & -2/3 & 7/6 \\ 1 & 0 & 1/2 \\ -1 & -3 & -5/2 \end{bmatrix}$

16) a) $[T]_{\beta_1\beta_2} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$ b) $[T]_{\beta_1'\beta_2'} = \begin{bmatrix} 3/2 & -3/2 \\ -3 & 6 \\ 1 & -4 \\ 0 & 1 \end{bmatrix}$

17) a) $[T]_\beta = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 3 & -2 \end{bmatrix}$ b) $[T]_{\beta'} = \begin{bmatrix} -25/4 & 13/4 & 10 \\ 7/4 & 5/4 & -2 \\ -25/8 & 5/8 & 5 \end{bmatrix}$

18) $[T]_\beta = \begin{bmatrix} 59/24 & 307/24 & -25/8 \\ 5/6 & 7/6 & 1/2 \\ 55/24 & 47/24 & 19/8 \end{bmatrix}$

SECCIÓN 4 (PÁGINAS 320-323)

1) $\beta' = \{(42, -5), (43, 5)\}$

2) $\beta' = \left\{\left(\frac{5}{2}, -\frac{3}{2}, -\frac{3}{2}\right), \left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right), \left(-\frac{3}{2}, \frac{3}{2}, \frac{5}{2}\right)\right\}$

3) Si A y B son semejantes, existe una matriz inversible P tal que $A = P^{-1}BP$. Entonces,

$$\begin{aligned} \det A &= \det(P^{-1}BP) \\ &= (\det P^{-1})(\det B)(\det P) \\ &= (\det P)^{-1}(\det B)(\det P) \\ &= \det B \end{aligned}$$

4) Si A y B fueran semejantes, existiría una matriz inversible de orden 2 $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tal que $PA = BP$, esto es, tal que $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ c & d \end{bmatrix}$ lo que obliga a que $c = d = 0$, contradiciendo la inversibilidad de P.

5) Al ser A y B semejantes, existe una matriz inversible P tal que $A = P^{-1}BP$. Entonces, $trA = tr(P^{-1}BP) = tr(P^{-1}PB) = trB$ (en donde se usó que $trMN = trNM$ —véase ejercicio 21 de la sección 3 del capítulo 1—.) El ejercicio anterior (4) muestra que la afirmación recíproca es falsa.

6) Según los ejercicios 3 y 5 anteriores, una condición necesaria para que dos matrices sean semejantes es que ambas tengan el mismo determinante o la misma traza. Entonces la matriz dada no es semejante a A puesto que:

a) no tiene el mismo determinante
b) no tiene el mismo determinante
c) no tiene la misma traza
d) no tiene el mismo determinante
e) no tiene el mismo determinante
f) no tiene el mismo determinante

7) Las matrices P, Q y R junto con sus inversas son:

$$P = \begin{bmatrix} -2 & 2 & 1 \\ 7/5 & -1 & -3/5 \\ 4/5 & -1 & -1/5 \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & -1 \\ 3 & 2 & 4 \end{bmatrix}$$

$$Q = \begin{bmatrix} 2/7 & 9/7 & -1 \\ 4/7 & -3/7 & 2 \\ 9/7 & 23/7 & -2 \end{bmatrix} \qquad Q^{-1} = \begin{bmatrix} -8 & -1 & 3 \\ 26/5 & 1 & -8/5 \\ 17/5 & 1 & -6/5 \end{bmatrix}$$

$$R = \begin{bmatrix} 3 & 2 & 0 \\ -1 & 0 & 1 \\ 0 & 3 & 1 \end{bmatrix} \qquad R^{-1} = \begin{bmatrix} 3/7 & 2/7 & -2/7 \\ -1/7 & -3/7 & 3/7 \\ 3/7 & 9/7 & -2/7 \end{bmatrix}$$

a) $[T]_{\beta_1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

b) $[T]_{\beta_2} = \begin{bmatrix} 8 & 4 & 13 \\ -16/5 & -2/5 & -34/5 \\ 17/5 & -9/5 & -28/5 \end{bmatrix}$

c) $[T]_{\beta_3} = \begin{bmatrix} 2 & 10/7 & 1/7 \\ -2 & -9/7 & 2/7 \\ 5 & 45/7 & 8/7 \end{bmatrix}$

8) Se tiene $B = PAP^{-1}$ con $P = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$

9) Si B está en la clase de equivalencia de A (esto es, si B es semejante a A) entonces (por el ejercicio 3) $\det B = \det A \neq 0$. Entonces, B es inversible.

10) a) Tiene sólo a la matriz cero.
b) Tiene sólo a la matriz identidad.
(Es decir que si A es semejante a la matriz cero —identidad—, entonces A *es* la matriz cero —identidad, respectivamente.)

11) Si A es semejante a B, existe una matriz inversible P tal que $A = P^{-1}BP$. Entonces, $A^2 = (P^{-1}BP)(P^{-1}BP) = P^{-1}B(PP^{-1})BP = P^{-1}B^2P$, de modo que A^2 es semejante a B^2. De manera análoga, A^n es semejante a B^n.

12) Existe $k \in \mathbf{N}$ tal que $A^k = 0$. Si B es semejante a A, entonces B^k es semejante a $A^k = 0$ (ejercicio anterior). Entonces, $B^k = 0$ [ejercicio 10, a).]

13) Si B es semejante a A, entonces B^2 es semejante a $A^2 = I$. Entonces, $B^2 = I$, esto es, B es involutoria.

14) Si B es semejante a A, entonces existe P inversible tal que $B = P^{-1}AP$, de donde $B^2 = P^{-1}A^2P = P^{-1}AP = B$, esto es, B es idempotente.

15) a) Las matrices $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$
son nilpotentes pero no son semejantes.

b) Las matrices $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
son involutorias pero no son semejantes.

c) Las matrices $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$
son idempotentes pero no son semejantes.

SECCIÓN 5 (PÁGINAS 331-343)

1) a) $(T_1 + T_2)(2, 1, 3) = (21, 15, 0)$
$(3T_1)(1, 1, 1) = (12, 3, 9)$
$(-2T_2)(2, 4, 0) = (-20, -16, 4)$

b) $(T_1 + T_2)(x, y, z) = (7x + y + 2z, x + y + 4z, 4x + y - 3z)$
$(cT_1)(x, y, z) = (c(2x + y + z), c(x - y + z), c(3x + 2y - 2z))$
$(cT_2)(x, y, z) = (c(5x + z), c(2y + 3z), c(x - y - z))$

3) a) $\beta = \{T_1, T_2\}$, $T_1, T_2 : \mathbf{R}^2 \to \mathbf{R}$
$T_1((x, y)) = x$, $T_2((x, y)) = y$

b) $\beta = \{T_1, T_2, T_3, T_4\}$, $T_1, \ldots, T_4 : \mathbf{R}^2 \to \mathbf{R}^2$
$T_1((x, y)) = (x, 0)$ $\quad T_3((x, y)) = (0, x)$
$T_2((x, y)) = (y, 0)$ $\quad T_4((x, y)) = (0, y)$

c) $\beta = \{T_1, \ldots, T_6\}$, $T_1, \ldots, T_6: \mathbf{R}^2 \to \mathbf{R}^3$

$T_1((x, y)) = (x, 0, 0)$ $\qquad$ $T_4((x, y)) = (0, y, 0)$

$T_2((x, y)) = (y, 0, 0)$ $\qquad$ $T_5((x, y)) = (0, 0, x)$

$T_3((x, y)) = (0, x, 0)$ $\qquad$ $T_6((x, y)) = (0, 0, y)$

d) $\beta = \{T_1, \ldots, T_6\}$, $T_1, \ldots, T_6: \mathbf{R}^3 \to \mathbf{R}^2$

$T_1((x, y, z)) = (x, 0)$ $\qquad$ $T_4((x, y, z)) = (0, x)$

$T_2((x, y, z)) = (y, 0)$ $\qquad$ $T_5((x, y, z)) = (0, y)$

$T_3((x, y, z)) = (z, 0)$ $\qquad$ $T_6((x, y, z)) = (0, z)$

e) $\beta = \{T_1, T_2, T_3, T_4\}$, $T_1, \ldots, T_4: P_1 \to P_1$

$T_1(a + bx) = a$ $\qquad$ $T_3(a + bx) = ax$

$T_2(a + bx) = b$ $\qquad$ $T_4(a + bx) = bx$

f) $\beta = \{T_1, \ldots, T_6\}$, $T_1, \ldots, T_6: P_1 \to P_2$

$T_1(a + bx) = a$ $\qquad$ $T_4(a + bx) = bx$

$T_2(a + bx) = b$ $\qquad$ $T_5(a + bx) = ax^2$

$T_3(a + bx) = ax$ $\qquad$ $T_6(a + bx) = bx^2$

g) $\beta = \{T_1, \ldots, T_{16}\}$, $T_1, \ldots, T_{16}: M_{2\times 2} \to M_{2\times 2}$

$T_1\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$ $\qquad$ $T_2\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} b & 0 \\ 0 & 0 \end{bmatrix}$

$T_3\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} c & 0 \\ 0 & 0 \end{bmatrix}$ $\qquad$ $T_4\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} d & 0 \\ 0 & 0 \end{bmatrix}$

$T_5\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}$ $\qquad$ $T_6\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}$

$T_7\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & c \\ 0 & 0 \end{bmatrix}$ $\qquad$ $T_8\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & d \\ 0 & 0 \end{bmatrix}$

$T_9\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ a & 0 \end{bmatrix}$ $\qquad$ $T_{10}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix}$

$T_{11}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix}$ $\qquad$ $T_{12}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ d & 0 \end{bmatrix}$

$T_{13}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & a \end{bmatrix}$ $\qquad$ $T_{14}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & b \end{bmatrix}$

$T_{15}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & c \end{bmatrix}$ $\qquad$ $T_{16}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & d \end{bmatrix}$

h) $\beta = \{T_1, T_2, T_3, T_4\}$, $T_1, \ldots, T_4: M_{1\times 2} \to M_{2\times 1}$

$T_1([a \quad b]) = \begin{bmatrix} a \\ 0 \end{bmatrix}$ $\qquad$ $T_2([a \quad b]) = \begin{bmatrix} b \\ 0 \end{bmatrix}$

$T_3([a \quad b]) = \begin{bmatrix} 0 \\ a \end{bmatrix}$ $\qquad$ $T_4([a \quad b]) = \begin{bmatrix} 0 \\ b \end{bmatrix}$

4) a) $\begin{bmatrix} 0 & 7 \\ 4 & 9 \\ -4 & 5 \end{bmatrix}$ $\qquad$ c) $\begin{bmatrix} -2 & 13 \\ 8 & 25 \\ -12 & 9 \end{bmatrix}$

b) $\begin{bmatrix} 2 & 8 \\ 4 & 2 \\ 0 & 6 \end{bmatrix}$ $\qquad$ d) $\begin{bmatrix} 7 & 14 \\ 6 & -11 \\ 8 & 11 \end{bmatrix}$

5) a) $[T_1 + T_2]_\beta = \begin{bmatrix} 25/7 & 55/7 \\ 214/63 & 17/7 \end{bmatrix}$

b) $[T_1 + T_2]_{\beta'} = \begin{bmatrix} 40/9 & 55/9 \\ 257/63 & 14/9 \end{bmatrix}$

c) $[T_1 + T_2]_{\beta_1} = \begin{bmatrix} -5/3 & 55/3 \\ 2/7 & 23/3 \end{bmatrix}$

d) $[T_1 + T_2]_{\beta_2} = \begin{bmatrix} 127/27 & -575/189 \\ -214/27 & 35/27 \end{bmatrix}$

e) $[T_1 + T_2]_{\beta_3} = \begin{bmatrix} 5585/189 & 7591/189 \\ -3190/189 & -4451/189 \end{bmatrix}$

f) $[T_1 + T_2]_{\beta_4} = \begin{bmatrix} 173/21 & 6/7 \\ -31/63 & -47/21 \end{bmatrix}$

7) a) $(T_1 \circ T_2)(x, y) = (7x + 2y, 14x + 8y)$

b) $(T_2 \circ T_1)(x, y) = (12x + y, 8x + 3y)$

c) $(T_1 \circ T_2 \circ T_3)(x, y) = (-9x - 2y, -22x - 8y)$

d) $(T_2 \circ (3T_1) \circ T_3)(x, y) = (-39x - 3y, -33x - 9y)$

e) $(T_2 \circ (T_1 + T_3))(x, y) = (5x - y, 3x + y)$

f) $(T_1 \circ (T_1 + 2T_2) \circ (3T_2 - T_3))(x, y) = (262x + 95y, 768x + 366y)$

14) a) $[T_1]_{\beta_1\beta_2} = \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 3 & -8 \end{bmatrix}$ $\qquad$ $[T_2]_{\beta_1\beta_2} = \begin{bmatrix} 1 & -1 & -1 \\ 2 & 5 & 1 \\ 3 & 8 & -2 \end{bmatrix}$

b) $(T_2 \circ T_1)(x, y) = (-2x + 4y, 12x + 5y, 8x + 37y)$

c) $\begin{bmatrix} -2 & 4 \\ 12 & 5 \\ 8 & 37 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ 2 & 5 & 1 \\ 3 & 8 & -2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 3 & -8 \end{bmatrix}$

15) $[T_1 \circ T_2]_\beta = \begin{bmatrix} 7 & 3 & 1 \\ 1 & 14 & 6 \\ 3 & 7 & 3 \end{bmatrix}$ $\qquad$ $[T_2 \circ T_1]_\beta = \begin{bmatrix} 1 & -3 & 4 \\ 8 & 10 & -2 \\ 11 & 5 & 13 \end{bmatrix}$

16) a) $[T_1 \circ T_2]_\beta = \begin{bmatrix} 4/3 & 413/42 \\ 6 & 18 \end{bmatrix}$, $[T_2 \circ T_1]_\beta = \begin{bmatrix} 215/21 & -565/42 \\ -200/21 & 191/21 \end{bmatrix}$

b) $[T_1 \circ T_2]_{\beta'} = \begin{bmatrix} 21 & 287/3 \\ 0 & -5/3 \end{bmatrix}$, $[T_2 \circ T_1]_{\beta'} = \begin{bmatrix} -8/3 & -71/3 \\ 1 & 22 \end{bmatrix}$

c) $[T_1 \circ T_2]_{\beta_1} = \begin{bmatrix} 18 & 6 \\ 413/42 & 4/3 \end{bmatrix}$, $[T_2 \circ T_1]_{\beta_1} = \begin{bmatrix} 191/21 & -200/21 \\ -565/42 & 215/21 \end{bmatrix}$

d) $[T_1 \circ T_2]_{\beta_2} = \begin{bmatrix} 21 & 413/6 \\ 0 & -5/3 \end{bmatrix}$, $[T_2 \circ T_1]_{\beta_2} = \begin{bmatrix} -50/3 & -565/6 \\ 6 & 36 \end{bmatrix}$

e) $[T_1 \circ T_2]_{\beta_3} = \begin{bmatrix} 131/6 & -5/6 \\ 47/2 & -5/2 \end{bmatrix}$, $[T_2 \circ T_1]_{\beta_3} = \begin{bmatrix} 203/18 & 11/6 \\ 11/6 & -13/2 \end{bmatrix}$

f) $[T_1 \circ T_2]_{\beta_4} = \begin{bmatrix} -5/2 & 47/2 \\ -5/6 & 131/6 \end{bmatrix}$, $[T_2 \circ T_1]_{\beta_4} = \begin{bmatrix} -13/2 & 11/6 \\ 11/6 & 203/18 \end{bmatrix}$

17) a) Tome $v \in V$, $v \in Ker\ T_1$. Entonces, $T_1(v) = 0$ y $T_2(T_1(v)) = T_2(0) = 0$, esto es $(T_2 \circ T_1)(v) = 0$, por lo que $v \in Ker\ T_2 \circ T_1$. Esto prueba que $Ker\ T_1 \subset Ker\ T_2 \circ T_1$.

b) Tome $w \in W$, $w \in Im T_2 \circ T_1$. Existe entonces $v \in V$ tal que $(T_2 \circ T_1)(v) = w$. Llame $u = T_1(v)$. Dado $w \in W$ se ha probado que existe $u = T_1(v) \in U$ tal que $T_2(u) = w$. Entonces, $w \in Im\ T_2$. Esto prueba que $Im T_2 \circ T_1 \subset Im T_2$.

18) Según el inciso a) del ejercicio anterior se tiene que nulidad $(T_2 \circ T_1) \geq$ nulidad T_1. Similarmente, el inciso b) del ejercicio anterior dice que rango $(T_2 \circ T_1) \leq$ rango T_2. Según el teorema de la dimensión se tiene

$$
\begin{aligned}
dim\ V &= \text{rango } T_1 + \text{nulidad } T_1 \\
&= \text{rango } T_2 + \text{nulidad } T_2 \\
&= \text{rango } (T_2 \circ T_1) + \text{nulidad } (T_2 \circ T_1)
\end{aligned}
$$

de donde: rango T_2 - rango $(T_2 \circ T_1)$ = nulidad $(T_2 \circ T_1)$ - nulidad $T_2 \geq 0$

nulidad $(T_2 \circ T_1)$ - nulidad T_1 = rango T_1 - rango $(T_2 \circ T_1) \geq 0$, por

lo que nulidad $(T_2 \circ T_1) \geq$ nulidad T_2

rango $T_1 \geq$ rango $(T_2 \circ T_1)$

que junto con nulidad $(T_2 \circ T_1) \geq$ nulidad T_1

rango $T_2 \geq$ rango $(T_2 \circ T_1)$ demuestra que

nulidad $(T_2 \circ T_1) \geq$ máx (nulidad T_1, nulidad T_2)

rango $(T_2 \circ T_1) \leq$ mín (rango T_1, rango T_2)

19) Suponga que se tiene la implicación $v \in Ker\ T^2 \Rightarrow v \in Ker\ T$, $v \in V$. Sea $v \in Ker\ T \cap ImT$. Como $v \in ImT$, existe $u \in V$ tal que $T(u) = v$, de donde $T^2(u) = T(v) = 0$, pues $v \in Ker\ T$. Entonces, $u \in Ker\ T^2$ y por tanto, $u \in Ker\ T$. Esto es, $v = T(u) = 0$. Esto muestra que $Ker\ T \cap ImT = \{0\}$. Suponga ahora que $Ker\ T \cap ImT = \{0\}$ y tome $v \in Ker\ T^2$. Se tiene $T(T(v)) = 0$, de donde $T(v) \in Ker\ T$. Pero obviamente $T(v) \in ImT$. Entonces, $T(v) \in Ker\ T \cap ImT$. Entonces, $T(v) = 0$, esto es, $v \in Ker\ T$. Entonces se ha probado que $v \in Ker\ T^2 \Rightarrow v \in Ker\ T$.

20) Según el teorema de la dimensión se tiene: $dim\ V$ = rango T + nulidad T = rango T^2 + nulidad T^2. Como rango T = rango T^2, se concluye que nulidad T = nulidad T^2. Según el inciso a) del ejercicio 17, se tiene que $Ker\ T \subset Ker\ T^2$. Como las dimensiones de estos dos espacios ($Ker\ T$ y $Ker\ T^2$) son iguales, se concluye que $Ker\ T = Ker\ T^2$. Por el ejercicio 19, se concluye, finalmente, que $Im\ T \cap Ker\ T = \{0\}$.

21) a) Tome $v \in Ker\ T$. Entonces, $v = v - T(v) = (Id - T)(v)$. Esto muestra que $v \in Im(Id - T)$. Entonces se ha probado la contención $Ker\ T \subset Im(Id - T)$ (observe que no se usa $T^2 = T$, esto es, esta contención es válida siempre). Tome ahora $v \in Im(Id - T)$. Existe entonces $u \in V$ tal que $u - T(u) = v$. Entonces $T(v) = T(u - T(u)) = T(u) - T^2(u) = T(u) - T(u) = 0$, esto es, $v \in Ker\ T$. Entonces se ha probado que $Im(Id - T) \subset Ker\ T$. Esto muestra el inciso a).

b) Tome $v \in Ker(Id - T)$. Entonces, $v - T(v) = 0$. Esto es, $v = T(v)$, lo que muestra que $v \in Im\ T$. Entonces se ha probado que $Ker(Id - T) \subset Im\ T$ (observe que esta contención es válida siempre). Tome ahora $v \in Im\ T$. Existe entonces $u \in V$ tal que $v = T(u)$. Entonces, $v - T(v) = T(u) - T(T(u)) = T(u) - T^2(u) = T(u) - T(u) = 0$, lo que muestra que $v \in Ker(Id - T)$. Entonces se ha probado que $Im\ T \in Ker(Id - T)$. Esto muestra el inciso b).

c) Se sigue inmediatamente del ejercicio 19), pues $Ker\ T^2 = Ker\ T$.

d) Cada vector $v \in V$ se puede escribir de manera única como $v = v - T(v) + T(v)$. Llame $u_1 = v - T(v)$ y $u_2 = T(v)$. Entonces, $v = u_1 + u_2$ en donde $T(u_1) = T(v - T(v)) = T(v) - T^2(v) = T(v) - T(v) = 0$, esto es, $u_1 \in Ker\ T$ y $u_2 = T(v) \in Im\ T$. Esto muestra $V = Ker\ T \oplus Im\ T$.

25) Es una aplicación directa de las propiedades 2) y 3) de la definición de álgebra dada en la página 354, pues $T(y_1 + y_2) = x_0(y_1 + y_2) = x_0y_1 + x_0y_2 = T(y_1) + T(y_2)$ y $T(cy) = x_0(cy) = c(x_0y) = cT(y)$.

26) a) Para cualquier $v \in A$ se tiene que $F(x + y)(v) = T_{x+y}(v) = (x + y)v = xv + yv = T_xv + T_yv = F(x)(v) + F(y)(v) = (F(x) + F(y))(v)$, de donde $F(x + y) = F(x) + F(y)$.

b) Para cualquier $v \in A$ se tiene que $F(cx)(v) = T_{cx}(v) = (cx)(v) = c(xv) = cT_xv = cF(x)(v)$, de donde $F(cx) = cF(x)$.

c) Para cualquier $v \in A$ se tiene que $F(xy)(v) = T_{xy}(v) = (xy)(v) = x(yv) = T_x(yv) = T_x(T_y(v)) = (T_xT_y)(v) = (F(x)F(y))(v)$, de donde $F(xy) = F(x)F(y)$.

d) Suponga que $F(x_1) = F(x_2) \in L(A, A)$. Entonces para cualquier $v \in A$ se tiene que $F(x_1)(v) = F(x_2)(v)$, o sea, $x_1v = x_2v$, de donde $(x_1 - x_2)v = 0$. En particular, si $v = e$ (el elemento unidad de A) se tiene $(x_1 - x_2)e = x_1 - x_2 = 0$, de donde $x_1 = x_2$. Esto muestra la inyectividad de F.

SECCIÓN 6 (PÁGINAS 348-358)

1) La afirmación recíproca es: Si $T_1: V \to U$ y $T_2: U \to W$ son transformaciones lineales tales que $T_2 \circ T_1: V \to W$ es inversible, entonces T_1 y T_2 son inversibles. Esta afirmación es falsa como lo muestra el siguiente ejemplo: Sea $T_1: \mathbf{R}^2 \to \mathbf{R}^3$ la transformación lineal $T(X) = AX$ en donde $A = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \\ 1/2 & 0 \end{bmatrix}$ y sea $T_2: \mathbf{R}^3 \to \mathbf{R}^2$ la transformación lineal $T_2(X) = BX$ en donde $B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

Entonces, $T_2 \circ T_1: \mathbf{R}^2 \to \mathbf{R}^2$ es la transformación $(T_2 \circ T_1)(X) = (BA)(X)$, en donde $BA = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \\ 1/2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Es decir, que $T_2 \circ T_1 = Id_{\mathbf{R}^2}$ que es obviamente inversible. Sin embargo, es claro que ni siquiera tiene sentido preguntarse por la inversibilidad de T_1 y T_2.

2) Sea $V = P$ = espacio de todos los polinomios en x. Sea $T: P_{n+1} \to P$ el operador $T(a_0 + a_1x + \ldots + a_nx^n) = a_0x + a_1\frac{x^2}{2} + \ldots + a_n\frac{x^{n+1}}{n+1}$, se afirma que T es inyectivo. En

efecto, si $p(x) = a_0 + a_1x + \ldots + a_nx^n \in Ker\ T$ entonces, $a_0x + a_1\frac{x^2}{2} + \ldots + a_n\frac{x^{n+1}}{n+1} = 0$, de donde $a_0 = a_1 = \ldots = a_n = 0$, esto es, $p = 0$. Entonces T es inyectivo. T no es sobreyectivo pues si $p \in P$ es un polinomio constante, $p \notin Im\ T$.

Considere ahora el operador $U: P \to P$, $U(a_0 + a_1x + \ldots + a_nx^n) = a_1 + \ldots + na_nx^{n-1}$. Se afirma que U no es inyectivo. En efecto, es claro que $Ker\ U$ = {polinomios constantes}. U sí es sobreyectivo. Dado cualquier $p \in P$, sea $\tilde{p} = Tp$ (en donde T es el operador dado en el ejemplo anterior). Entonces, $U\tilde{p} = p$, esto es, $p \in Im\ U = P$.

3) a) Se tiene que $Ker\ (T_1 \circ T_2) = \{0\}$. Según el ejercicio 17 de la sección 5, se tiene que $Ker\ T_2 \subseteq Ker\ (T_1 \circ T_2) = \{0\}$. Entonces, $Ker\ T_2 = \{0\}$, esto es, T_2 es inyectivo.

b) Se tiene que $Im\ (T_1 \circ T_2) = V$. Según el ejercicio 17 de la sección anterior se tiene que $V = Im\ (T_1 \circ T_2) \subseteq Im\ T_1$. Entonces, $Im\ T_1 = V$, esto es, T_1 es sobreyectivo.

c) y d) No. Considere por ejemplo los operadores lineales $T_1 = U$ y $T_2 = T$ en el espacio vectorial P dados en el ejercicio anterior. Se tiene que $T_1 \circ T_2 = U \circ T = Id_P$ que es inyectivo y sobreyectivo. Sin embargo, se vio que $T_2 = T$ no es sobreyectivo y $T_2 = U$ no es inyectivo.

4) Es una consecuencia inmediata del teorema 6.6, pues $det\ T = ad - bc$.

5) Según el ejercicio 3, se tiene que siendo $S \circ T = Id_V$ inyectivo y sobreyectivo, entonces T es inyectivo. Como $dim\ V = dim\ U$ (finita), el teorema 6.4 permite concluir que T es inversible.

6) Para la transformación lineal $T_1: \mathbf{R}^3 \to \mathbf{R}^2$ se tiene que rango $T_1 \leq 2$.

Por el ejercicio 18 de la sección anterior se tiene rango $(T_2 \circ T_1) \leq$ mín (rango T_1, rango $T_2) \leq 2$. Entonces $T_2 \circ T_1: \mathbf{R}^3 \to \mathbf{R}^3$ no puede ser sobreyectiva. Por el teorema 6.2 se concluye que $T_2 \circ T_1$ no puede ser inversible.

7) Para la transformación lineal $T_1: \mathbf{R}^n \to \mathbf{R}^m$ se tiene que rango $T_1 \leq m$.

Por el ejercicio 18 de la sección anterior se tiene rango $(T_2 \circ T_1) \leq$ mín (rango T_1, rango $T_2) \leq m < n$. Entonces, $T_2 \circ T_1: \mathbf{R}^n \to \mathbf{R}^n$ no puede ser sobreyectiva. Por el teorema 6.2 se concluye que $T_2 \circ T_1$ no puede ser inversible.

9) a) La contención $Ker\ T_2 \subseteq Ker\ (T_1 \circ T_2)$ siempre es cierta (véase ejercicio 17 de la sección anterior). Sea entonces $v \in Ker\ (T_1 \circ T_2)$, esto es, $(T_1 \circ T_2)(v) = 0$ o $T_1(T_2v) = 0$. Entonces, $T_2(v) \in Ker\ T_1 = \{0\}$, esto es, $T_2(v) = 0$. Por lo tanto, $v \in Ker\ T_2$. Esto muestra que $Ker\ (T_1 \circ T_2) \subseteq Ker\ T_2$, y por tanto, que $Ker\ T_2 = Ker\ (T_1 \circ T_2)$.

b) La contención $Im\ (T_1 \circ T_2) \subseteq Im\ T_1$ siempre es cierta (véase ejercicio 17 de la sección anterior). Sea entonces $v \in Im\ T_1$. Existe $u \in V$ tal que $T_1(u) = v$. Pero $V = Im\ T_2$, por lo que se puede escribir $u = T_2(w)$ para algún vector $w \in V$. Entonces, $T_1(T_2(w)) = v$, esto es, $(T_1 \circ T_2)(w) = v$, lo que muestra que $v \in Im(T_1 \circ T_2)$. Se ha probado así que $Im\ T_1 \subseteq Im\ (T_1 \circ T_2)$ y por tanto, que $Im\ T_1 = Im(T_1 \circ T_2)$.

10) Se tiene $Ker\ D$ = {polinomios constantes}. Entonces, D no es inyectivo y por lo tanto, no es inversible.

11) Tome $v \in Ker\ (Id - T)$. Entonces, $v - T(v) = 0$, de donde $T(v - T(v)) = T(0) = 0$, o sea, $0 = T(v) - T^2(v) = T(v) - 0 = T(v)$. Por lo tanto, $v = 0$, esto es, $Ker\ (Id - T) = \{0\}$.

12) a) $T^{-1}: \mathbf{R}^2 \to \mathbf{R}^2$, $T^{-1}(x, y) = \left(\frac{1}{11}x + \frac{2}{11}y, -\frac{3}{11}x + \frac{5}{11}y\right)$

b) $T^{-1}: \mathbf{R}^3 \to \mathbf{R}^3$, $T^{-1}(x, y, z) = \left(\frac{1}{2}x - \frac{2}{5}y + \frac{11}{45}z, \frac{1}{5}y + \frac{2}{45}z, \frac{1}{9}z\right)$

c) $T^{-1}: P_2 \to \mathbf{R}^3$, $T^{-1}(a + bx + cx^2) = (a, -a + b, a - b + c)$

d) $T^{-1}: M_{2\times2} \to M_{2\times2}$, $T^{-1}(A) = A'$

e) $T^{-1}: M_{2\times2} \to M_{2\times2}$, $T^{-1}(A) = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} A$

f) $T: P_3 \to M_{2\times2}$, $T^{-1}(a + bx + cx^2 + dx^3) = \begin{bmatrix} d & c - d \\ b - c & a - b \end{bmatrix}$

13) T es inversible pues $det\ T = sgn\begin{pmatrix} 1 & 2 & \ldots & n \\ n & n-1 & \ldots & 1 \end{pmatrix} \neq 0$

Se tiene de hecho $T^{-1} = T$

14) a) T no es inyectiva puesto que $I \in Ker\ T$. En efecto, $T(I) = IB - BI = B - B = 0$. Por el teorema 6.2, T no es inversible.

b) $[T]_\beta = \begin{bmatrix} -1 & 2 & 1 & 2 \\ 4 & 4 & 4 & 0 \\ -3 & 0 & -1 & 2 \\ 0 & -3 & -2 & -2 \end{bmatrix}$

Se tiene $det\ [T]_\beta = 0$, por lo que T no es inversible (teorema 6.6).

15) T es inversible si, y sólo si $Ker\ T = \{0\}$, esto es, si, y sólo si se tiene la implicación $BA = 0 \Rightarrow A = 0$. Pero esta implicación es cierta si, y sólo si B es una matriz inversible. El operador inverso es $T^{-1}(A) = B^{-1}A$.

16) a) $det\ T = -92$. El operador sí es inversible.

b) $det\ T = 0$. El operador no es inversible.

c) $det\ T = 1$. El operador sí es inversible.

d) $det\ T = 0$. El operador no es inversible.

e) $det\ T = 0$. El operador no es inversible.

f) $det\ T = 0$. El operador no es inversible.

SECCIÓN 7 (PÁGINAS 366-374)

1) Si $\lambda = 3$ se tiene rango = 2. Si $\lambda \neq 3$ se tiene rango = 3.

3) Si $\lambda = 0$, la matriz tiene rango 2. Este el mínimo rango que se puede tener.

4) a) Si $a = 1$, rango = 2. Si $a = -14$, rango = 3.

b) Si $a = 0$, rango = 2. Si $a = 1$ o $a = \frac{21}{20}$, rango = 3.

5) Si $a + b = 0$, el rango es 2.

Si $b = 3a$, el rango es 3.

6) $t = 0, -1$.

7) Una posible demostración es como sigue: Sean $T_1, T_2: \mathbf{R}^n \to \mathbf{R}^n$ las transformaciones lineales $T_1X = AX$, $T_2X = BX$. Se afirma que $Im(T_1 + T_2) \subseteq Im\ T_1 + Im\ T_2$. En efecto, tome $y \in Im(T_1 + T_2)$. Existe entonces $x \in \mathbf{R}^n$ tal que $(T_1 + T_2)(X) = T_1X + T_2X = y$, esto es, $y \in Im\ T_1 + Im\ T_2$. Por lo tanto, se tiene $dim\ Im\ (T_1 + T_2) \leq dim\ (Im\ T_1 + Im$

$T_2) = dim\ Im\ T_1 + dim\ Im\ T_2 - dim\ ((Im\ T_1) \cap (Im\ T_2)) \leq dim\ Im\ T_1 + dim\ Im\ T_2$ o bien, en términos matriciales (teorema 7.2) rango $(A + B) \leq$ rango A + rango B.

8) Si T_1, T_2: $\mathbf{R}^n \to \mathbf{R}^n$ son las transformaciones lineales $T_1X = BX$, $T_2X = AX$, se sabe (por el ejercicio 18 de la sección 5)

$$\text{rango } (T_1 \circ T_2) \leq mín \text{ (rango } T_2, \text{ rango } T_1).$$

El escribir esta desigualdad en términos matriciales (teorema 7.2) se obtiene rango $(AB) \leq mín$ (rango A, rango B).

9) La matriz que representa al operador T respecto de la base canónica β de $M_{n \times n}$ es la matriz de orden n^2 siguiente:

$$[T]_\beta = \begin{bmatrix} B^t & & 0 \\ & B^t & \\ 0 & \ddots & B^t \end{bmatrix}$$

Entonces, el rango de T se puede hallar como el número de líneas no nulas en la forma escalonada reducida de la matriz

$$C = \begin{bmatrix} B & & 0 \\ & B & \\ 0 & \ddots & B \end{bmatrix}$$

(véase teorema 7.1). Llevando B a la forma escalonada reducida se obtiene rango B líneas no nulas. Haciendo las operaciones elementales correspondientes a este proceso en las primeras n líneas de C, luego en las segundas n líneas de C, etc., se obtiene finalmente una matriz con n(rango B) líneas no nulas. Éste es, entonces, el rango de T.

10) a) Llame $A' = BA$. Usando el mismo argumento de la demostración del teorema 7.1, capítulo 3, se puede demostrar (siendo B inversible) que el espacio línea de A es el mismo que el espacio línea de A', esto es, de BA. De aquí se concluye, entonces, que el rango de A es igual al rango de BA.

b) Por el inciso anterior se tiene que rango A = rango B_1A. Para ver que rango B_1A = rango B_1AB_2, use el inciso b) del ejercicio 9 de la sección anterior, identificando al operador T_1 con la matriz B_1A y al operador T_2 con la matriz B_2 (que es inversible).

12) a) 0; b) 1; c) 1; d) 2; e) 0; f) 2; g) 3; h) 2; i) 3.

13) De denotará por r_m al rango de la matriz del sistema y por r_{ma} al rango de la matriz aumentada del sistema.

f) $r_m = r_{ma} = 1$. El sistema es consistente.

a), b), d) $r_m = r_{ma} = 2$. El sistema es constante.

c), g), h), j) $r_m = r_{ma} = 3$. El sistema es consistente.

i) $r_m = r_{ma} = 4$. El sistema es consistente.

o) $r = 2, r = 3$. El sistema no es consistente.

APÉNDICE I (PÁGINAS 376-387)

1) Se tiene:

$$f(A + B) = tr(A + B) = tr((a_{ij} + b_{ij})_{i,j=1,\ldots,n}) = \sum_{i=1}^{n} (a_{ii} + b_{ii})$$

$$= \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = tr\,A + tr\,B = f(A) + f(B)$$

$$\text{si } c \in \mathbf{R}, f(cA) = tr(cA) = tr((ca_{ij})_{i,j=1,\ldots,n}) = \sum_{i=1}^{n} ca_{ii} = c\sum_{i=1}^{n} a_{ii}$$

$c\,tr\,A = cf(A)$

2) Se tiene $\varphi(f_1 + f_2) = (f_1 + f_2)(0) = f_1(0) + f_2(0) = \varphi(f_1) + \varphi(f_2)$

Si $c \in \mathbf{R}$, $\varphi(cf) = (cf)(0) = cf(0) = c\varphi(f)$

3) $f(x, y) = \frac{4}{3}x - \frac{1}{3}y$

4) Falso. La afirmación es cierta si, y sólo si, v_1 y v_2 son linealmente independientes, esto es, si $\{v_1, v_2\}$ es una base de $\mathbf{R}^2$.

5) $f(x, y, z) = \frac{1}{23}(5x + 19y + 11z)$

6) La base dual de β es $\{f_1, f_2, f_3\}$ en donde f_i: $\mathbf{R}^3 \to \mathbf{R}$, $i = 1, 2, 3$, son los funcionales lineales

$f_1(x, y, z) = 3y - x$

$f_2(x, y, z) = x - 2y$

$f_3(x, y, z) = \frac{1}{4}(7y - 3x + z)$

7) La base dual de β es $\{f_1, f_2, f_3, f_4\}$ en donde f_i: $M_{2 \times 2} \to \mathbf{R}$, $i = 1, 2, 3, 4$ son los funcionales lineales

$$f_1\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \frac{1}{5}(d - c)$$

$$f_2\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = b - \frac{1}{5}c - \frac{4}{5}d$$

$$f_3\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \frac{1}{2}a - \frac{3}{10}c - \frac{1}{5}d$$

$$f_4\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = -\frac{1}{2}a + \frac{11}{10}c + \frac{2}{5}d$$

8) $T^t f$: $\mathbf{R}^2 \to \mathbf{R}$ está dada por

a) $(T^t f)(x_1, x_2) = ax_1 + bx_2$

b) $(T^t f)(x_1, x_2) = bx_2$

c) $(T^t f)(x_1, x_2) = (a + b)x_1 + (a - b)x_2$

d) $(T^t f)(x_1, x_2) = (3a - 2b)x_1 + (4a + 6b)x_2$

9) El funcional $T^t f$: $M_{n \times n} \to \mathbf{R}$ es $(T^t f)(A) = tr(BA)$

10) a) $(1, 1) = 0(2, 3) + (1, 1)$

b) $(4, 6) = 2(2, 3) + (0, 0)$

c) $(5, 5) = 0(2, 3) + (5, 5)$

d) $(-1, 2) = 3(2, 3) + (-7, -7)$

11) $(3, 1, 2) = -2(-1, -1, -1) + (1, -1, 0)$

12) $f: \mathbf{R}^3 \to \mathbf{R}$ es el funcional lineal $f(x, y, z) = 2x - y + z$

13) b) $f: \mathbf{R}^3 \to \mathbf{R}$ es el funcional lineal $f(x, y, z) = x + 4y - 2z$

14) b) $f: M_{2\times 2} \to \mathbf{R}$ es el funcional lineal $f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = d$

15) El funcional lineal $f: \mathbf{R}^4 \to \mathbf{R}$ es $f(x_1, x_2, x_3, x_4) = (2x_4 - 2x_2 + x_3 - x_1)c$ en donde c es un número (arbitrario pero fijo) no nulo. Se tiene, entonces, $M = \{v \in \mathbf{R}^4 \mid f(v) = c\}$ o sea $(0, 1, 1, 1) + \mathscr{L}((1, 0, 1, 0), (0, 1, 0, 1), (2, 0, 0, 1)) = \{(x_1, x_2, x_3, x_4) \in \mathbf{R}^4 \mid 2x_4 - 2x_2 + x_3 - x_1 = 1\}$.

APÉNDICE II (PÁGINAS 389-400)

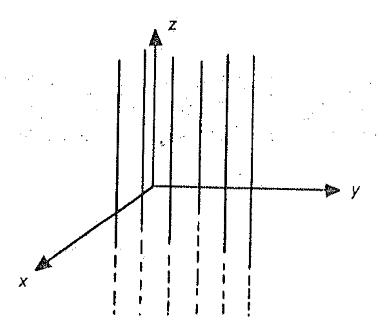
1) En el espacio V/V sólo existe la clase de equivalencia del vector 0. En el espacio $V/\{0\}$ se tienen tantas clases de equivalencia como vectores en V; se tiene de hecho un isomorfismo entre V y $V/\{0\}$.

2) a) Son planos paralelos al plano xy.

b) Una base de $\mathbf{R}^3/W$ está constituida por el único vector (clase de equivalencia) $(0, 0, 1) + W$.

3) a) Una recta del tipo $x = x_0$, $y = y_0$, $z = t$, $t \in \mathbf{R}$.

b) Son rectas en $\mathbf{R}^3$ paralelas al eje z.



c) Una base de $\mathbf{R}^3/W$ es $\{(1, 0, 0) + W, (0, 1, 0) + W\}$

4) Una base de P/P_4 está constituida por el conjunto (infinito) de clases de equivalencia $\{x^5 + P_4, x^6 + P_4, \ldots, x^n + P_4, \ldots\}$

5) El isomorfismo es $\varphi: V/U \to \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a = c = d = 0\right\}$

$$\varphi\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + U\right) = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}$$

CAPÍTULO CINCO

SECCIÓN 1 (PÁGINAS 404-411)

1) Como $(v_0 \mid v) = 0 \ \forall v \in V$, en particular $(v_0 \mid v_0) = 0$, lo cual implica que $v_0 = 0$.

2) Si $(\cdot \mid \cdot)_1$ y $(\cdot \mid \cdot)_2$ son dos productos internos en V, la función $(\cdot \mid \cdot): V \times V \to \mathbf{R}$, $(\cdot \mid \cdot) = (\cdot \mid \cdot)_1 - (\cdot \mid \cdot)_2$ no es necesariamente un producto interno en V, pues, por ejemplo, no es cierto en general que $(v \mid v) \geq 0 \ \forall v \in V$.

3) a) Sí es un producto interno.

b) No es un producto interno: no es cierto que $(x \mid x) \geq 0 \ \forall x \in \mathbf{R}^2$, por ejemplo si $x = (0, 1)$, se tiene $(x \mid x) = -1$.

c) Sí es un producto interno.

d) Sí es un producto interno.

e) No es un producto interno: no es lineal.

4) a) Sí es un producto interno.

b) Sí es un producto interno.

c) No es un producto interno: no es cierto que $(x \mid x) \geq 0 \ \forall x \in \mathbf{R}^3$, por ejemplo si $x = (1, 3, 1)$, $(x \mid x) = -8$.

d) No es un producto interno: no es cierto que $(x \mid y) = (y \mid x) \ \forall x, y \in \mathbf{R}^3$.

e) No es un producto interno: no es cierto que $(x \mid y) = (y \mid x) \ \forall x, y \in \mathbf{R}^3$.

5) a) Sí es un producto interno.

b) Sí es un producto interno.

c) Sí es un producto interno.

d) No es un producto interno.

6) a) No es un producto interno: no es lineal.

b) No es un producto interno: no es cierto que $(A \mid A) \geq 0 \ \forall A \in M_{n\times n}$, por ejemplo, si $n = 2$ y $A = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$, tenemos $(A \mid A) = -2$.

c) Sí es un producto interno.

8) Sí es un producto interno en P_1. No es un conjunto interno en P_3: no es cierto que $(p \mid p) = 0 \Rightarrow p = 0$. Por ejemplo, si $p(x) = x(x - 1)(x - 2) \in P_3$ se tiene $(p \mid p) = 0$. Tampoco es un producto interno en P_n, $n \geq 4$.

9) $u = (c_1, c_2 - c_1, c_3 - c_2)$.

10) $A = \begin{bmatrix} 1/9 & 4/9 \\ 8/3 & 14/9 \end{bmatrix}$

11) Cualquier vector $(x, y, z) \in \mathbf{R}^3$ tal que $x + 5y + 4z = 0$ cumple la condición dada.

12) Al poner $v_1 = (a_{11}, a_{12}, a_{13})$, $v_2 = (a_{21}, a_{22}, a_{23})$ se tiene que $(x \mid v_1) = (x \mid v_2) = 0$ se ve como $a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = 0$ $a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = 0$ que es un sistema homogéneo de 2 ecuaciones lineales con 3 incógnitas, el cual se sabe que tiene una infinidad de soluciones.

15) Es la desigualdad de Cauchy-Schwartz para los vectores $x = (x_1, x_2, \ldots, x_n)$ y $y = (y_1, y_2, \ldots, y_n)$ en $\mathbf{R}^n$ con el producto interno introducido en la página 453.

16) Suponga primeramente que los vectores v y u en el espacio vectorial V son linealmente dependientes. Existe entonces $k \in \mathbf{R}$ tal que $u = kv$, y $(v \mid u)^2 = (v \mid kv)^2 = k^2(v \mid v)^2 = (v \mid v)(kv \mid kv) = (v \mid v)(u \mid u)$, teniéndose así la igualdad en la desigualdad de Cauchy-Schwartz. Suponga ahora que los vectores $u, v \in V$ son tales que $(v \mid u)^2 = (v \mid v)(u \mid u)$ (esto es, que se cumple la igualdad en la desigualdad de Cauchy-Schwartz). Si $v = 0$, no hay nada que hacer pues los vectores v y u serían linealmente dependientes (teorema 4.2 (2) del capítulo 3). Si $v \neq 0$, la función cuadrática $f(t) = at^2 + bt + c$, con $a = (v \mid v)$, $b = 2(u \mid u)$, $c = (u \mid u)$ tendría una raíz doble, dígase

$\tilde{t}(= -(v \mid u)/(v \mid v))$. Entonces, $(\tilde{t}v + u \mid \tilde{t}u + u) = 0$ de donde $\tilde{t}v + u = 0$, lo que muestra la dependencia lineal de u y v.

SECCIÓN 2 (PÁGINAS 414-419)

1) a) 1 b) 1 c) 1 d) $\sqrt{26}$ e) 1 f) $\sqrt{10}$

2) a) 1 b) $\sqrt{2}$ c) $\sqrt{2}$ d) $\sqrt{59}$ e) 1 f) $\sqrt{21}$

3) a) 1 b) $\sqrt{2}$ c) $\sqrt{30}$ d) $\sqrt{18}$ e) 2 f) $\sqrt{14}$

4) a) $\sqrt{\operatorname{sen} h\, 2}$ b) ≈ 0.7385 c) $\sqrt{2/3}$ d) $\sqrt{2/5}$ e) ≈ 1.206 f) ≈ 1.3912

6) $\|-v\| = \|(-1)v\| = |-1|\ \|v\| = 1 \cdot \|v\| = \|v\|$

7) a) 6 b) 6 c) 9 d) 24

9) a) $\|v\| = \sqrt{26}$, $\|u\| = \sqrt{234}$, $\|v + u\| = \sqrt{416}$
b) $\|v + u\| = \sqrt{416} = \sqrt{26} + \sqrt{234} = \|v\| + \|u\|$

10) $\|v\| = \sqrt{107}$, $\|u\| = \sqrt{963}$, $\|v + u\| = \sqrt{1712}$
Sí se cumple la igualdad en la desigualdad triangular.

11) Suponga primero que $u = kv$, $k \geq 0$. Entonces, $\|u + v\| = \|kv + v\| = |k + 1|\ \|v\| = (k + 1)\|v\| = k\|v\| + \|v\| = \|kv\| + \|v\| = \|u\| + \|v\|$, teniéndose así la igualdad deseada. Suponga ahora que se tienen dos vectores u y v en el espacio vectorial V para los cuales $\|u + v\| = \|u\| + \|v\|$ (esto es, para los que se cumple la igualdad en la desigualdad triangular). Se tiene $\|u + v\|^2 = (\|u\| + \|v\|)^2 = \|u\|^2 + 2\|u\|\ \|v\| + \|v\|^2$, pero $\|u + v\|^2 = (u + v \mid u + v) = \|u\|^2 + 2(u \mid v) + \|v\|^2$. Por lo que $(u \mid v) = \|u\|\ \|v\|$, es decir, se tiene la igualdad en la desigualdad de Cauchy-Schwartz. Por el ejercicio 16 de la sección anterior, existe $k \in \mathbf{R}$ tal que $u = kv$. Sustituyendo este vector en la igualdad $\|u + v\| = \|u\| + \|v\|$, se obtiene $|k + 1|\ \|v\| = (|k| + 1)\|v\|$. Si $v = 0$, el resultado se sigue de inmediato. Si $v \neq 0$, entonces $|k + 1| = |k| + 1$, lo que implica que $k \geq 0$.

12) Se tiene $\|v\| = \|v - u + u\| \leq \|v - u\| + \|u\|$ de donde $\|v - u\| \geq \|v\| - \|u\|$. También $\|u\| = \|u - v + v\| \leq \|v - u\| + \|v\|$, de donde $-\|v - u\| \leq \|v\| - \|u\|$. Entonces, $|\ \|v\| - \|u\|\ | \leq \|v - u\|$

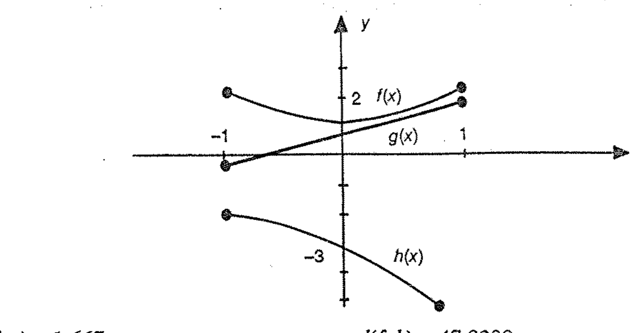
13) Se tiene $\|v + u\|^2 + \|v - u\|^2 = (v + u \mid v + u) + (v - u \mid v - u) = \|v\|^2 + 2(v \mid u) + \|u\|^2 + \|v\|^2 - 2(v \mid u) + \|u\|^2 = 2(\|v\|^2 + \|u\|^2)$

14) a) $\sqrt{5}$ b) $\sqrt{2}$ c) $\sqrt{20}$

15) a) $\sqrt{30}$ b) $\sqrt{45}$ c) $\sqrt{17}$

16) $d(f, g) = 4.2776 < 4.6776 = d(f, h)$. Entonces la función g está más cerca de f que h.

17)



$d(f, g) \approx 1.667$ $d(f, h) \approx 47.3239$

Entonces g se encuentra más cerca de f que h.

20) d) d.1) $\|\cdot\|_1 = 1$, $\|\cdot\|_2 = 1$
d.2) $\|\cdot\|_1 = 11$, $\|\cdot\|_2 = 5$
d.3) $\|\cdot\|_1 = 4$, $\|\cdot\|_2 = 1$
d.4) $\|\cdot\|_1 = 31$, $\|\cdot\|_2 = 20$
d.5) $\|\cdot\|_1 = 3$, $\|\cdot\|_2 = 3$
d.6) $\|\cdot\|_1 = 11$, $\|\cdot\|_2 = 6$
d.7) $\|\cdot\|_1 = 14$, $\|\cdot\|_2 = 8$
d.8) $\|\cdot\|_1 = 5$, $\|\cdot\|_2 = 3$

e) e.1) 2
e.2) 1
e.3) 1
e.4) sen 1
e.5) e
e.6) 2
e.7) 8
e.8) 5

f) Es la misma demostración del ejercicio 13 de esta sección.

g) Si $v = (1, 0)$, $u = (0, 1)$ se tiene $\|v + u\|_1 = \|v - u\|_1 = 2$,
$\|v\|_1 = \|u\|_1 = 1$.
No se cumple la ley del paralelogramo.

h) Si $v = (1, 0)$, $u = (0, 1)$ se tiene $\|v + u\|_2 = \|v - u\|_2 = \|v\|_2 = \|u\|_2 = 1$.
No se cumple la ley del paralelogramo.

i) Si $f(x) = x^2$, $g(x) = x$ se tiene $\|f\| = \|g\| = 1$, $\|f + g\| = \|f - g\| = 2$
No se cumple la ley del paralelogramo.

21) f) No se cumple la desigualdad triangular, por ejemplo con $x = -1$, $z = 1$, $y = 0$.

i) Si $x \neq y$, $c \neq 0$ se tiene $d(cx, cy) = 1 = d(x, y)$.
(No se cumple la segunda condición establecida en el inciso h)).

SECCIÓN 3 (PÁGINAS 429-439)

1) a) *arccos* $1/\sqrt{65}$ b) *arccos* $3/\sqrt{10}$ c) *arccos* $(-4/\sqrt{65})$ d) $\pi/2$ *rad* e) *arccos* $(-4/5)$ f) π *rad*

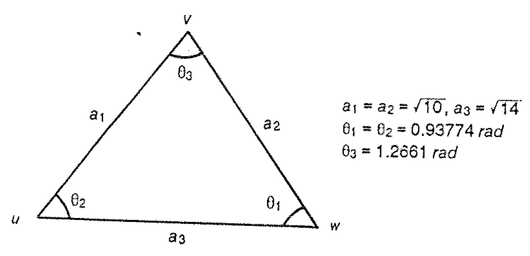
2) a) *arccos* $4/\sqrt{66}$ b) $\pi/2$ *rad* c) *arccos* $13/15\,\sqrt{2}$

3) a) $arccos\, 0.60234 \approx 0.92437\; rad$ c) $arccos\, 0.3528 \approx 1.2102\; rad$

b) $\pi/2\; rad$

4) $\pi\; rad$

5)

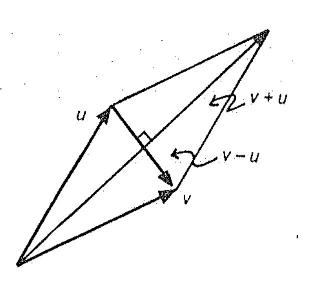


6) Los vectores del inciso d) del ejercicio 1.
Los vectores del inciso b) del ejercicio 2.
Los vectores del inciso b) del ejercicio 3.

7) b) $arccos\,(-13/3\sqrt{65})$

8) Véase ejercicio 1 de la sección 1 de este capítulo.

9) $(v + u \mid v - u) = \|v\|^2 - (v \mid u) + (u \mid v) - \|u\|^2 = 0$ pues $\|v\| = \|u\|$



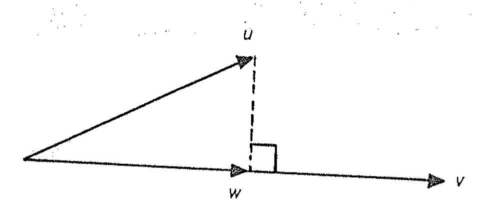
10) Es una consecuencia directa del corolario del teorema 3.2 de este capítulo y del inciso 4 del teorema 5.5 del capítulo 3.

13) a) No pertenece a $\mathcal{L}(S)$ c) Sí pertenece a $\mathcal{L}(S)$

b) Sí pertenece a $\mathcal{L}(S)$

14) $C_1 = \dfrac{1}{\sqrt{2}}(x_1 + x_3),\; C_2 = \dfrac{1}{\sqrt{2}}(-x_2 + x_4),\; C_3 = \dfrac{1}{2}(x_1 - x_2 - x_3 - x_4)$

15) a)



b) $\|w\| = \left\| \dfrac{(u \mid v)}{\|v\|^2}\, v \right\| = |(u \mid v)|$, pues $\|v\| = 1$.

c) El vector cero.

d) Sea $(V, (\cdot \mid \cdot))$ un espacio con producto interno y sea W el subespacio de V generado por el conjunto ortonormal $S = \{v_1, v_2, \ldots, v_k\}$. Si u es un vector de W, entonces el cuadrado de la norma de u es igual a la suma de los cuadrados de las normas de las proyecciones ortogonales de u sobre cada uno de los vectores $v_1, v_2, \ldots, v_k$.

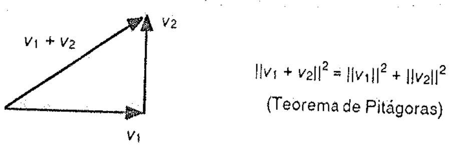
e) El conjunto ortonormal $\{v_1, v_2\}$ es entonces una base de $\mathbb{R}^2$. Según el teorema 3.3 todo vector $\mathbb{R}^2$ se puede escribir como $u = (u \mid v_1)v_1 + (u \mid v_2)v_2$. Pero $w_i = (u \mid v_i)$ v_i = proyección ortogonal de u sobre v_i, $i = 1, 2$.

16) $u = \sum_{i=1}^{k} v_i$

$$\|u\|^2 = (u \mid u) = \left(\sum_{i=1}^{k} v_i \;\middle|\; \sum_{j=1}^{k} v_j \right)$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{k} (v_i \mid v_j) = \sum_{i=1}^{k} (v_i \mid v_i) = \sum_{i=1}^{k} \|v_i\|^2$$

Por ejemplo, si $k = 2$

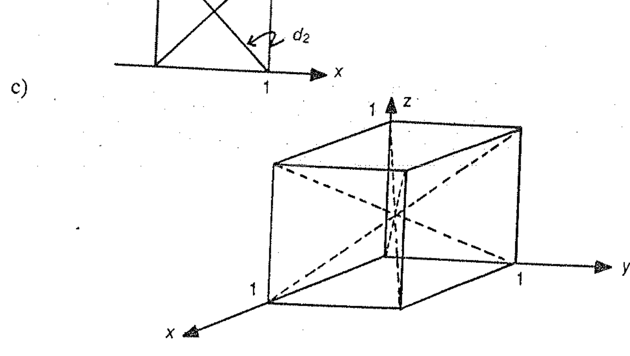


17) a)

0 1

b)

c)



d) 2^{n-1} diagonales distintas.

e) Si n es impar, no existen diagonales ortogonales. Si n es *par*, dígase $n = 2k$, el número de diagonales ortogonales a una de ellas es $\frac{1}{2}\binom{n}{k}$.

f) $\sqrt{n}$

h) h.1) Para $n = 2$

h.2) Para $n = 4$

SECCIÓN 4 (PÁGINAS 443-461)

1) a) $\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} = 5\begin{bmatrix} ½ & ½ \\ ½ & ½ \end{bmatrix} - 2\begin{bmatrix} ½ & ½ \\ -½ & -½ \end{bmatrix} + \begin{bmatrix} ½ & -½ \\ -½ & -½ \end{bmatrix}$

b) $\begin{bmatrix} -1 & 3 \\ 2 & 1 \end{bmatrix} = \frac{5}{2}\begin{bmatrix} ½ & ½ \\ ½ & ½ \end{bmatrix} - \frac{1}{2}\begin{bmatrix} ½ & ½ \\ -½ & -½ \end{bmatrix} - \frac{3}{2}\begin{bmatrix} ½ & -½ \\ ½ & -½ \end{bmatrix} - \frac{5}{2}\begin{bmatrix} ½ & -½ \\ -½ & ½ \end{bmatrix}$

c) $(v \mid w) = (5)\left(\frac{5}{2}\right) + (-2)\left(-\frac{1}{2}\right) + (0)\left(-\frac{3}{2}\right) + (1)\left(-\frac{5}{2}\right) = 11$

d) $\|v\| = \sqrt{(5)^2 + (-2)^2 + (1)^2} = \sqrt{30}$

$\|u\| = \sqrt{\left(\frac{5}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + \left(-\frac{3}{2}\right)^2 + \left(-\frac{5}{2}\right)^2} = \sqrt{15}$

2) $u_1 = \frac{1}{\sqrt{2}}(1, 1), \quad u_2 = \frac{1}{\sqrt{2}}(1, -1)$

3) $u_1 = \frac{1}{2}(1, 1, 1), \quad u_2 = \frac{1}{2\sqrt{43}}(1, 5, -11) \quad u_3 = \frac{1}{\sqrt{86}}(8, -3, -2)$

4) $u_1 = 1,\ u_2 = x,\ u_3 = x^2,\ u_4 = x^3$

5) a) $u_1 = \frac{1}{\sqrt{5}}(1, 0, 2), \quad u_2 = \frac{1}{\sqrt{70}}(-6, 5, 3)$

b) $u_2 = \frac{1}{\sqrt{14}}(3, 1, -2)$

c) $u_1 = (1, 0, 0), \quad u_2 = (0, 1, 0)$

d) $u_1 = \frac{1}{\sqrt{3}}(1, 1, 1)$

Nota: la respuesta en los incisos a) y c) no es única.

6) a) $u_1 = (1, 0, 0, 0)$

b) $u_1 = \frac{1}{\sqrt{2}}(1, 0, 1, 0), \quad u_2 = \frac{1}{\sqrt{2}}(0, 1, 0, 1)$

c) $u_1 = \frac{1}{\sqrt{2}}(1, 0, 0, -1), \quad u_2 = \frac{1}{\sqrt{6}}(-1, 2, 0, -1), \quad u_3 = \frac{1}{\sqrt{12}}(-1, -1, 3, -1)$

d) $u_1 = \frac{1}{2}(1, 1, 1, 1)$

e) $u_1 = \frac{1}{\sqrt{15}}(1, 2, 3, 1), \quad u_2 = \frac{1}{\sqrt{85}}(6, -3, -2, 6)$

f) $u_1 = \frac{1}{\sqrt{18}}(2, 3, -1, 2)$

g) $u_1 = \frac{1}{\sqrt{2}}(1, 0, 1, 0), \quad u_2 = \frac{1}{\sqrt{2}}(0, 1, 0, 1), \quad u_3 = \frac{1}{\sqrt{2}}(-1, 0, 1, 0)$

h) $u_1 = (1, 0, 0, 0),\ u_2 = (0, 1, 0, 0),\ u_3 = (0, 0, 1, 0),\ u_4 = (0, 0, 0, 1)$

7) a) $u_1 = \frac{1}{\sqrt{2}}(0, 1, -1)$

b) $u_1 = \frac{1}{\sqrt{62}}(1, 5, 6)$

c) $u_1 = \frac{1}{\sqrt{13}}(-2, 3, 0, 0, 0)$

$u_2 = \frac{1}{3\sqrt{27326}}(-144, -96, 351, -234, 195)$

Nota: La respuesta en el inciso c) no es única.

8) $p(x) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}x, \quad q(x) = x^2, \quad r(x) = \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}x$

9) $u_1 = \frac{1}{\sqrt{3}}(1, 1, 1)$

$u_2 = \frac{1}{\sqrt{6}}(-1, 2, -1)$

$u_3 = \frac{1}{\sqrt{2}}(-1, 0, 1)$

(La respuesta no es única.)

10) $u_1 = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \end{bmatrix} \qquad u_3 = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & -1/\sqrt{2} \end{bmatrix}$

$u_2 = \begin{bmatrix} 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 \end{bmatrix} \qquad u_4 = \begin{bmatrix} 0 & 1/\sqrt{2} \\ -1/\sqrt{2} & 0 \end{bmatrix}$

(La respuesta no es única.)

11) $u_1 = \frac{1}{\sqrt{2}}, \quad u_2 = \sqrt{\frac{3}{2}}\,x, \quad u_3 = \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right), \quad u_4 = \sqrt{\frac{175}{8}}\left(x^3 - \frac{3}{5}x\right)$

$1 + 2x - 3x^2 + x^3 = \frac{26}{15}\sqrt{\frac{3}{2}}\,u_2 - \sqrt{\frac{8}{5}}\,u_3 + \sqrt{\frac{8}{175}}\,u_4$

12) La j-ésima columna de la matriz $[T]_\beta$ está constituida por las coordenadas del vector $T(v_j) \in V$ respecto de la base (ortonormal) β. El resultado se sigue entonces del inciso 2) del teorema 4.1.

13) Siendo A ortogonal se tiene $A \cdot A^t = I$, lo cual se puede escribir como $A^tA = I$, o sea, $A^t(A^t)^t = I$. De aquí que $(A^t)^{-1} = (A^t)^t$, o sea que A^t también es ortogonal.

14) Al usar el resultado del ejercicio 14 de la sección 2, capítulo 2, se tiene:
$(AB)(AB)^t = (AB)(B^tA^t) = A(BB^t)A^t = AA^t = I$
$(BA)(BA)^t = (BA)(A^tB^t) = B(AA^t)B^t = BB^t = I$ es decir que $(AB)^{-1} = (AB)^t$ y
$(BA)^{-1} = (BA)^t$, o sea, AB y BA son ortogonales.

15) Si A es ortogonal, $AA^t = I$. Si A es simétrica $A = A^t$. Entonces, $A \cdot A = I$, esto es, $A^2 = I$.

16) Escriba $A = (a_{ij})_{i,j=1,\ldots,n}$. Observe que A es ortogonal si, y sólo si $AA^t = I$ o $A^tA = I$, es decir, si, y sólo si

$$\sum_{k=1}^{n} a_{ik}a_{jk} = \delta_{ij} = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

(Lo que significa que los vectores línea de A forman una base ortonormal de $\mathbf{R}^n$), o

$$\sum_{k=1}^{n} a_{ki}a_{kj} = \delta_{ij} = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

(Lo que significa que los vectores columna de A forman una base ortonormal de $\mathbf{R}^n$), respectivamente.

18) $\alpha = \beta = 0$

21) Suponga que $u \in W^{\perp}$, entonces $(u \mid v) = 0 \ \forall\, v \in W$. En particular para los vectores $v_1, v_2, \ldots, v_k \in W$ se tiene $(u \mid v_i) = 0$, $i = 1, 2, \ldots, k$. Suponga ahora que $(u \mid v_i) = 0$, $i = 1, 2, \ldots, k$. Sea v un vector cualquiera de W, existen entonces escalares $c_1, c_2, \ldots, c_k$ tales que $v = c_1v_1 + c_2v_2 + \ldots + c_kv_k$. Entonces, $(u \mid v) = (u \mid c_1v_1 + c_2v_2 + \ldots + c_kv_k) = c_1(u \mid v_1) + c_2(u \mid v_2) + \ldots + c_k(u \mid v_k) = 0$, esto es, $(u \mid v) = 0 \ \forall\, v \in W$, o bien, $u \in W^{\perp}$.

23) $x_1 = 2t, x_2 = -3t, x_3 = t, t \in \mathbf{R}$

24) Falso. La afirmación es cierta si, y sólo si los vectores v_1 y v_2 son linealmente independientes.

25) $W^{\perp}$ es la recta que pasa por el origen $x_1 = at, x_2 = bt, x_3 = ct, t \in \mathbf{R}$.

26) a) $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2} \mid a = d = 0\right\}$

b) $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2} \mid a = b = c = d\right\}$

c) $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2} \mid b = c = d = 0\right\}$

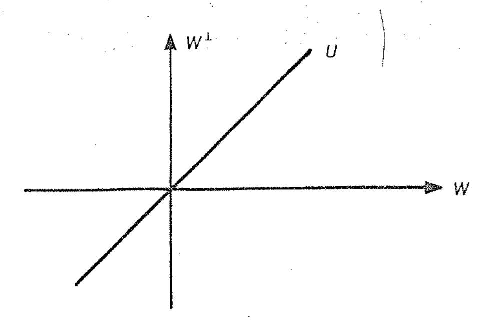
d) $\{a \in M_{2\times 2} \mid A = -A^t\}$

e) $\{a \in M_{2\times 2} \mid A = A^t\}$

27) $\{a_0 + a_1x + a_2x^2 + a_3x^3 \mid 3a_0 + a_2 = 0\}$

28) $W^{\perp} = \mathcal{L}((2, -1, 0, -8), (0, 0, 1, 3))$.

29) No. Por ejemplo, $V = \mathbf{R}^2$ con el producto interno canónico. $W = \mathcal{L}((1, 0))$, $U = \mathcal{L}((1, 1))$. Se tiene $\mathbf{R}^2 = W \oplus U$, pero $U \neq W^{\perp}$.



32) Sea $u \in W_2^{\perp}$. Entonces $(u \mid v) = 0 \ \forall\, v \in W_2$. En particular, como $W_1 \subset W_2$, se tiene $(u \mid v) = 0 \ \forall\, v \in W_1$. Entonces, $u \in W_1^{\perp}$, esto es, $W_2^{\perp} \subset W_1^{\perp}$.

33) La contención $W \subset (W^{\perp})^{\perp}$ es inmediata: tome $w \in W$ y u un vector cualquiera en $W^{\perp}$. Entonces, $(w \mid u) = 0$, lo que significa que $w \in (W^{\perp})^{\perp}$. Pruebe entonces que $(W^{\perp})^{\perp} \subset W$. Sea x un vector cualquiera de $(W^{\perp})^{\perp}$ (subespacio del espacio vectorial V). Por el teorema 4.5 se puede escribir $x = w_1 + w_2$, en donde $w_1 \in W$ y $w_2 \in W^{\perp}$, es decir, que $w_2 = x - w_1$, en donde $x \in (W^{\perp})^{\perp}$ y $w_1 \in W \subset (W^{\perp})^{\perp}$. Como $(W^{\perp})^{\perp}$ es un subespacio de V, se concluye que $w_2 \in (W^{\perp})^{\perp}$. O sea que $w_2 \in W^{\perp} \cap (W^{\perp})^{\perp}$. Vea que $w_2 = 0$: como $w_2 \in (W^{\perp})^{\perp}$, $(w_2 \mid w) = 0 \ \forall\, w \in W^{\perp}$; como $w_2 \in W^{\perp}$, $(w_2 \mid w_2) = 0$, de donde $w_2 = 0$. Entonces, $x = w_1 \in W$, como se quería probar.

34) a) $v \in (W_1 + W_2)^{\perp} \Leftrightarrow (v \mid w_1 + w_2) = 0 \ \forall\, w_1 \in W_1, \forall\, w_2 \in W_2 \Leftrightarrow (v \mid w_1) = 0 \ \forall\, w_1 \in W_1$ y $(v \mid w_2) = 0 \ \forall\, w_2 \in W_2 \Leftrightarrow v \in W_1^{\perp} \cap W_2^{\perp}$

b) La igualdad mostrada en el inciso a) aplicada a los subespacios $W_1^{\perp}$ y $W_2^{\perp}$ da $(W_1^{\perp} + W_2^{\perp})^{\perp} = (W_1^{\perp})^{\perp} \cap (W_2^{\perp})^{\perp} = W_1 \cap W_2$ (por el resultado del ejercicio anterior). Es decir, que $W_1 \cap W_2 = (W_1^{\perp} + W_2^{\perp})^{\perp}$, de donde $(W_1 \cap W_2)^{\perp} = [(W_1^{\perp} + W_2^{\perp})^{\perp}]^{\perp} = W_1^{\perp} + W_2^{\perp}$.

36) a) $(1, 3, 1) = \underbrace{(0, 1, 2)}_{\in W} + \underbrace{(1, 2, -1)}_{\in W^{\perp}}$ c) $(3, 1, 2) = \underbrace{(2, 2, 2)}_{\in W} + \underbrace{(1, -1, 0)}_{\in W^{\perp}}$

b) $(2, 1, 0) = \underbrace{(0, 0, 0)}_{\in W} + \underbrace{(2, 1, 0)}_{\in W^{\perp}}$ d) $(1, 1, -2) = \underbrace{(1, 1, -2)}_{\in W} + \underbrace{(0, 0, 0)}_{\in W^{\perp}}$

37) a) $(1, 3, 1, 1) = \underbrace{(2, 2, 1, 1)}_{\in W} + \underbrace{(-1, 1, 0, 0)}_{\in W^{\perp}}$

b) $(2, 1, 1, 0) = \underbrace{(1, 0, 0, -1)}_{\in W} + \underbrace{(1, 1, 1, 1)}_{\in W^{\perp}}$

c) $(1, 3, 1, 0) = \underbrace{(7/12, 35/12, 11/12, 3/4)}_{\in W} + \underbrace{(5/12, 1/12, 1/12, -3/4)}_{\in W^{\perp}}$

c) Sí es un vector propio con $\lambda = 3$.
d) Sí es un vector propio con $\lambda = 2$.
e) No es un vector propio.
f) Sí es un vector propio con $\lambda = 1$.
i) Sí es un vector propio con $\lambda = 2$.
j) Sí es un vector propio con $\lambda = 3$.
k) No es un vector propio.
l) No es un vector propio.

4) El polinomio característico de A es $p(x) = (\lambda_1 - x)(\lambda_2 - x) \ldots (\lambda_n - x)$. Los valores propios son $\lambda_1, \lambda_2, \ldots, \lambda_n$.

5) Si $A\ (a_{ij})_{i,j=1,\ldots,n}$ es triangular, el polinomio característico de A es $p(x) = (a_{11} - x)(a_{22} - x) \ldots (a_{nn} - x)$. Los valores propios son: $a_{11}, a_{22}, \ldots, a_{nn}$.

6) El polinomio característico de Id es $p(x) = (1 - x)^n$, en donde $n = dim\ V$. El único valor propio es $\lambda = 1$. Cualquier vector no nulo de V es un vector propio de Id.

7) El polinomio característico de 0 es $p(x) = (-1)^n x^n$, en donde $n = dim\ V$. El único valor propio es $\lambda = 0$. Cualquier vector no nulo de V es un vector propio de 0.

8) Polinomio característico de $A = det(A - \lambda I) = det(A - \lambda I)^t = det(A^t - \lambda I)$ = polinomio característico de A^t. En general no tienen los mismos vectores propios.

9) Por inducción. Si $n = 1$ se trata de la hipótesis del ejercicio, suponga válido el resultado para $n = k$, esto es, suponga que $A^k X = \lambda^k X$, X un vector no nulo de $\mathbb{R}^n$, n = orden de A. Entonces, $A(A^k X) = A(\lambda^k X)$, o sea, $A^{k+1} X = \lambda^k AX = \lambda^k \lambda X = \lambda^{k+1} X$, lo que prueba que el resultado es válido para $n = k + 1$.

10) Sea X un vector propio asociado al valor propio λ, entonces $AX = \lambda X$. Multiplicando por la izquierda por A^{-1}, se obtiene $X = \lambda A^{-1} X$, de donde $A^{-1} X = \lambda^{-1} X$, lo que muestra que λ^{-1} es un valor propio de A^{-1}. Se observa que A y A^{-1} tienen los mismos vectores propios.

11) Sea X un vector propio asociado al valor propio λ, entonces $AX = \lambda X$, de donde $cAX = c\lambda X$ y así, $c\lambda$ es un valor propio de cA.

12) El polinomio característico de A es un polinomio de grado n (impar), el cual tiene al menos una raíz real.

13) No, en general.

14) Sea X un vector propio asociado al valor propio λ. Entonces, $(A + cI)X = AX + cX = (\lambda + c)X$, lo que muestra que $\lambda + c$ es un valor propio de $A + cI$.

15) Al usar el ejercicio 9 se demuestra fácilmente que el único valor propio de la matriz nilpotente A es $\lambda = 0$.

17) El polinomio característico de A es $p(\lambda) = det(A - \lambda I)$. Su término independiente se obtiene haciendo $\lambda = 0$. Entonces éste es $p(0) = det(A - 0 \cdot I) = det\ A$.

18) Se tiene A inversible $\Leftrightarrow det\ A \neq 0 \Leftrightarrow$ el término independiente del polinomio característico de A es $\neq 0 \Leftrightarrow \lambda = 0$ no es una raíz del polinomio característico de $A \Leftrightarrow \lambda = 0$ no es un valor propio de A.

19) a) Se tiene $(I - BA)(I + B(I - AB)^{-1} A) = I + B(I - AB)^{-1} A - BAB(I - AB)^{-1} - BA$ $= I + B((I - AB)^{-1} - AB(I - AB)^{-1} - I)\,A = I + B((I - AB)(I - AB)^{-1} - I)A = I + B(I - I)A = I$ de donde $(I - BA)^{-1} = I + B(I - AB)^{-1}A$.

b) Si $\lambda = 0$ es un valor propio de AB entonces, $0 = p(\lambda) = det\,(AB) = det\,(BA)$ lo que muestra que $\lambda = 0$ es también un valor propio de BA. Suponga ahora que $\lambda \neq 0$ es un valor propio de AB. Entonces, $0 = p(\lambda) = det\,(AB - \lambda I) = (-\lambda)^n\, det\,(I - \frac{1}{\lambda} AB)$, esto es, la matriz $I - \frac{1}{\lambda} AB$ no es inversible. Por el inciso anterior la matriz $I - \frac{1}{\lambda}$

3) base de $E_{\lambda_1} = \{(1, x^2 + x^3)\}$
base de $E_{\lambda_2} = \{-2x^2 + x^3\}$

h) 1) $p(\lambda) = (\lambda - 1)^4$
2) $\lambda = 1$
3) base de $E_\lambda = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right\}$

SECCIÓN 2 (PÁGINAS 530-534)

1) En cada caso se puede ver que el polinomio característico de la matriz no se puede factorizar como producto de factores lineales

a) $p(\lambda) = \lambda^2 - 2\lambda + 9$
b) $p(\lambda) = \lambda^2 - 3\lambda + 8$
c) $p(\lambda) = \lambda^2 - 5\lambda + 38$
d) $p(\lambda) = (1 - \lambda)(\lambda^2 - 3\lambda + 12)$
e) $p(\lambda) = (\lambda - 1)^2 (\lambda^2 + \lambda + 19)$

2) En el ejercicio 23 de la sección anterior, inciso c), se vio que esta matriz tiene sólo un valor propio ($\lambda = 1$) y que la dimensión del espacio propio correspondiente es 1. La conclusión se sigue entonces del teorema 2.3.

3) La matriz A tiene 4 valores propios distintos $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3, \lambda_4 = 4$. Por el corolario del teorema 2.2, A es diagonalizable. La matriz P es:

$$P = \begin{bmatrix} 1 & 2 & 9 & 32 \\ 0 & 1 & 6 & 24 \\ 0 & 0 & 2 & 12 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

5) a) Sí es diagonizable.
b) Posiblemente sea diagonizable, sí $dim\ E_{\lambda=1} = 2$ y $dim\ E_{\lambda=-1} = 2$.
c) Posiblemente sea diagonizable, sí $dim\ E_{\lambda=0} = 5$.
d) No es diagonalizable.
e) Sí es diagonalizable.
f) No es diagonalizable.
g) No es diagonalizable.
h) Sí es diagonalizable.
i) No es diagonalizable.
j) Posiblemente sea diagonalizable, si $dim\ E_{\lambda=\sqrt{2}} = dim\ E_{\lambda=-\sqrt{2}} = 2$.

6) Siendo A diagonalizable, A es semejante a la matriz diagonal $D = diag\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Entonces $tr\ A = tr\ D = \lambda_1 + \lambda_2 + \ldots + \lambda_n$ y $det\ A = det\ D = \lambda_1 \lambda_2 \ldots \lambda_n$.

8) Sí es diagonalizable. Si se considera la base de $M_{2\times 2}$

$$\beta = \left\{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$$

d) $(2, 1, 0, 1) = \underbrace{(37/17, 12/17, 1/17, 13/17)}_{\in W} + \underbrace{(-3/17, 5/17, -1/17, 4/17)}_{\in W^{\perp}}$

38) $(7, -4, -1, 2) = \underbrace{(5, -6, -2, -1)}_{\in W} + \underbrace{(2, 1, 1, 3)}_{\in W^{\perp}}$

39) $y = 2.3913x + 0.4316$

40) a) $pH = 9.0$, $T = 40°C$, $k = 0.00998727$
b) $pH = 9.0$, $T = 50°C$, $k = 0.014264$
c) $pH = 11.0$, $T = 40°C$, $k = 0.07254$
d) $pH = 11.0$, $T = 50°C$, $k = 0.214076$

41) $y = 4.7292x^2 - 2.3235x + 1.7727$

SECCIÓN 5 (PÁGINAS 485-488)

6) El vector $(a, b) \in \mathbf{R}^2$ puede ser $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ o $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$

7) La matriz que representa a T respecto de la base β puede ser

$$\begin{bmatrix} 2/3 & 1/\sqrt{2} & 1/\sqrt{18} \\ 1/3 & 0 & -4/\sqrt{18} \\ 2/3 & -1/\sqrt{2} & 1/\sqrt{18} \end{bmatrix} \text{ o } \begin{bmatrix} 2/3 & 1/\sqrt{2} & -1/\sqrt{18} \\ 1/3 & 0 & 4/\sqrt{18} \\ 2/3 & -1/\sqrt{2} & -1/\sqrt{18} \end{bmatrix}$$

8) a) Sea θ_1 el ángulo entre el vector u y el vector v, y θ_2 el ángulo entre $T(u)$ y $T(v)$ ($0 \leq \theta_1, \theta_2 \leq \pi$). Entonces, $\cos\theta_2 = \frac{(T(u)|T(v))}{\|T(u)\|\,\|T(v)\|} = \cos\theta_1$ de donde $\theta_2 = \theta_1$, esto es, T es conforme.

b) La transformación $T: V \to V$, $Tv = 2v$ es conforme pues

$$\frac{(T(u)|T(v))}{\|T(u)\|\,\|T(v)\|} = \frac{(2u\,|\,2v)}{\|2u\|\,\|2v\|} = \frac{4(u\,|\,v)}{4\|u\|\,\|v\|} = \frac{(u\,|\,v)}{\|u\|\,\|v\|}$$

Sin embargo, T no es ortogonal, pues $\|T(v)\| = \|2v\| \neq \|v\|$.

9) a) Una rotación de 90°.
b) Una rotación de 45°.
c) Una rotación de 60°.
d) Una rotación de 30°.

11) La composición de dos rotaciones es una rotación.

12) No. Si T_1 y T_2 son dos transformaciones tales que $det\ T_1$ y $det\ T_2$ es 1 o -1, no es cierto, en general, que $det\ (T_1 + T_2)$ sea 1 o -1.

13) a) Una reflexión respecto del eje x seguida de una rotación de 90°.
b) Una reflexión respecto del eje x seguida de una rotación de 45°.
c) Una reflexión respecto del eje x seguida de una rotación de 60°.
d) Una reflexión respecto del eje x seguida de una rotación de 30°.

15) Sea $A = (a_{ij})_{i,j=1,\ldots,n}$. Siendo A antisimétrica se tiene $a_{ij} = -a_{ji}$ (y entonces, $a_{ii} = 0$).

a) Sea $x = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$. Se tiene $(Ax\,|\,x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_ix_j = 0$.

b) $\|(A - I)\,x\|^2 = (Ax - x\,|\,Ax - x) = (Ax\,|\,Ax) - (Ax\,|\,x) - (x\,|\,Ax) + (x\,|\,x) = \|Ax\|^2 + \|x\|^2$ (usando el resultado de a)).

c) Suponga que $(A - I)x = 0$, entonces $\|(A - I)x\|^2 = \|Ax\|^2 + \|x\|^2 = 0$, de donde $\|x\| = 0$, esto es, $x = 0$. Esto prueba que $A - I$ es inversible.

d) $\|(A + I)x\|^2 = (Ax + x\,|\,Ax + x) = \|Ax\|^2 + \|x\|^2 = \|(A - I)\,x\,\|^2$.

e) Aplique el resultado del inciso anterior al vector $(A - I)^{-1}x$ para obtener $\|(A + I)(A - I)^{-1}x\|^2 = \|x\|^2$. Esto muestra que T es ortogonal y por tanto, isométrica.

f) $U - I = (A + I)(A - I)^{-1} - (A - I)(A - I)^{-1} = (A + I - A + I)(A - I)^{-1} = 2(A - I)^{-1}$.

APÉNDICE (PÁGINAS 492-493)

2) Si $f_1, f_2 \in V^*$ y $c \in \mathbf{R}$ entonces, $T(f_1 + cf_2) = \tilde{u}$ en donde $(f_1 + cf_2)(v) = (v\,|\,\tilde{u})$, $u \in V$, esto es, $f_1(v) + cf_2(v) = (v\,|\,\tilde{u}) = (v\,|\,u_{f_1}) + c(v\,|\,u_{f_2}) = (v\,|\,u_{f_1} + cu_{f_2})$ de donde $(v\,|\,\tilde{u} - (u_{f_1} + cu_{f_2})) = 0$. Como el vector $v \in V$ fue arbitrario, se concluye que $\tilde{u} = u_{f_1} + cu_{f_2}$, esto es, $T(f_1 + cf_2) = T(f_1) + cT(f_2)$, lo que muestra que T es lineal. Suponga ahora que $f \in Ker\ T$, esto es, que $u_f = 0$. Entonces, $f(v) = (v\,|\,u_f) = 0\ \forall\ v \in V$, de donde $f = 0$. Se ha probado así que $Ker\ T = \{0\}$. Entonces, T es un isomorfismo.

3) La afirmación completa sería: "se dice que el funcional $f \in V^*$ es ortogonal a $v_0 \in V$ si el vector $u_f \in V$ (asociado a f según el isomorfismo del ejercicio anterior) es ortogonal a v_0".

4) Primeramente observe que

$$W_1^{\perp} = \{f \in V^*\,|\,f(v) = 0\ \forall\ v \in W_1\} \quad W_2^{\perp} = \{f \in V^*\,|\,f(v) = 0\ \forall\ v \in W_2\}$$

son subespacios de V^*. Sea $\beta_1 = \{v_1, v_2, \ldots, v_k\}$ una base de W_1 y $\beta_2 = \{u_1, u_2, \ldots, u_m\}$ una base de W_2. Entonces $\beta = \{v_1, \ldots, v_k, u_1, \ldots, u_m\}$ es una base de V (pues $V = W_1 \oplus W_2$). Considere la base dual $\beta' = \{f_1, \ldots, f_k, g_1, \ldots, g_m\}$ de la base β. Cada $f \in V^*$ se puede escribir *de manera única* como $f = c_1f_1 + \ldots + c_kf_k + d_1g_1 + \ldots + d_mg_m$. ($c_i, d_j \in \mathbf{R}$, $i = 1, \ldots, k$, $j = 1, \ldots, m$). Escriba $\varphi = d_1g_1 + \ldots + d_mg_m$ y $\psi = c_1f_1 + \ldots + c_kf_k$, y observe que $\varphi \in W_1^{\perp}$ y $\psi \in W_2^{\perp}$ (por la manera como se define la base dual). Entonces, $f = \varphi + \psi$, $\varphi \in W_1^{\perp}$, $\psi \in W_2^{\perp}$. Esto concluye la prueba de que $V^* = W_1^{\perp} \oplus W_2^{\perp}$.

CAPÍTULO SEIS

SECCIÓN 1 (PÁGINAS 507-518)

1) a) No es un vector propio.
b) Sí es un vector propio con $\lambda = 4$.
c) Sí es un vector propio con $\lambda = 4$.
d) Sí es un vector propio con $\lambda = 0$.
e) Sí es un vector propio con $\lambda = 0$.
f) No es un vector propio.

2) El vector $X \in \mathbf{R}^n$ es no nulo y $AX = 0 = 0 \cdot X$

3) a) No es un vector propio.
b) Sí es un vector propio con $\lambda = 2$.
g) Sí es un vector propio con $\lambda = 3$.
h) Sí es un vector propio con $\lambda = 1$.

La matriz que representa a T respecto de β es

$$[T]_\beta = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(Véase ejercicio 24 de la sección anterior, inciso f).)

9) a) $\beta = \{(0, 1, -1), (-1, 4, 3), (0, 1, 1)\}$ $\quad [T]_\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ $\quad \det T = 6$

b) $\beta = \{(1, 2, 1), (1, 1, 0), (1, 2, 2)\}$ $\quad [T]_\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ $\quad \det T = 6$

c) $\beta = \{(2, 1, 0), (-1, 0, 1), (3, 5, 6)\}$ $\quad [T]_\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ $\quad \det T = -1$

d) $\beta = \{(1, 0, 1), (3, 2, 1), (1, 3, 1)\}$ $\quad [T]_\beta = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ $\quad \det T = -2$

e) $\beta = \{(1, 3, 0), (0, -2, 1), (2, 1, 2)\}$ $\quad [T]_\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ $\quad \det T = 2$

10) a) $\beta = \left\{ \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 5 & 5 \\ 2 & 4 \end{bmatrix} \right\}$

$$[T]_\beta = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \qquad \det T = -36$$

b) $\beta = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -3 & -4 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix} \right\}$.

$$[T]_\beta = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \qquad \det T = 0$$

c) $\beta = \left\{ \begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

$$[T]_\beta = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \det T = -2$$

11) a) $A^n = I$ si n es par, $A^n = A$ si n es impar.

b) $A^n = \begin{bmatrix} -2 + 3 \cdot 2^n & -3 + 3 \cdot 2^n \\ 2 - 2^{n+1} & 3 - 2^{n+1} \end{bmatrix}$

c) $A^n = \begin{bmatrix} \frac{1}{3}((-1)^{n+1} + 2^{n+2}) & \frac{4}{3}((-1)^n - 2^n) \\ \frac{1}{3}((-1)^{n+1} + 2^n) & \frac{1}{3}((-1)^n 4 - 2^n) \end{bmatrix}$

d) $A^n = \begin{bmatrix} (-1)^n + 2^{n+2} & (-1)^{n+1} 4 - 2^{n+3} & (-1)^{n+1} \cdot 2 - 2^{n+1} \\ (-1)^n + 2^{n+1} & (-1)^{n+1} 4 - 2^{n+2} & (-1)^{n+1} \cdot 2 - 2^n \\ (-1)^{n+1} 2 - 2^{n+1} & (-1)^n \cdot 8 + 2^{n+2} & (-1)^n \cdot 4 + 2^n \end{bmatrix}$

e) $A^n = \begin{bmatrix} -2^n & 2^{n+2} & 3 \cdot 2^n \\ -2^{n+1} & 5 \cdot 2^n & 3 \cdot 2^n \\ 2^{n+1} & -2^{n+2} & -2^{n+1} \end{bmatrix}$

14) a) $\sqrt{A_1} = \begin{bmatrix} 3 & 3/2 \\ 2/3 & 3 \end{bmatrix}, \sqrt{A_2} = -\sqrt{A_1}, \sqrt{A_3} = \begin{bmatrix} 1 & 9/2 \\ 2 & 1 \end{bmatrix}, \sqrt{A_4} = -\sqrt{A_3}$

b) Existen 8 posibles respuestas.

$$\sqrt{A_1} = \begin{bmatrix} 2 & -10 & -10 \\ 0 & -5 & -7 \\ 0 & 2 & 4 \end{bmatrix}, \sqrt{A_2} = -\sqrt{A_1}, \sqrt{A_3} = \begin{bmatrix} 2 & -10 & -10 \\ 0 & -17/5 & -7/5 \\ 0 & 2/5 & -8/5 \end{bmatrix}, \sqrt{A_4} = -\sqrt{A_3},$$

$$\sqrt{A_5} = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 17/5 & 7/5 \\ 0 & -2/5 & 8/5 \end{bmatrix}, \sqrt{A_6} = -\sqrt{A_5}, \sqrt{A_7} = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 5 & 7 \\ 0 & -2 & -4 \end{bmatrix}, \sqrt{A_8} = -\sqrt{A_7}$$

SECCIÓN 3 (PÁGINAS 548-551)

1) El polinomio mínimo de la matriz identidad es $m(x) = x - 1$ y el de la matriz cero es $m(x) = x$.

2) Polinomio característico: $p(\lambda) = (c - \lambda)^n$. Polinomio mínimo: $m(x) = x - c$.

5) (Se denotará por $p(x)$ al polinomio característico y por $m(x)$ al polinomio mínimo.)

a) $p(x) = x^2 + 11x + 30 = m(x)$

b) $p(x) = (x - 1)^2 = m(x)$

c) $p(x) = x^2 - x - 1 = m(x)$

d) $p(x) = (1 - x)(2 - x)(3 - x) = -m(x)$

e) $p(x) = -x^3 + 4x^2 - 4x + 1 = -m(x)$

f) $p(x) = -(x - 1)^3 = -m(x)$

g) $p(x) = (2 - x)^4 = m(x)$

h) $p(x) = (2 - x)^4, m(x) = (x - 2)^3$

i) $p(x) = (2 - x)^4, m(x) = (x - 2)^2$

6) Por el corolario del teorema 3.4, el polinomio característico de A debe ser $p(x) = (a - x)\, m(x)$, lo cual contradice al teorema 3.5 (el factor $(a - x)$ que aparece en $p(x)$ debería aparecer en $m(x)$).

7) El polinomio característico de A es $p(x) = (-1)^n x^n$. Por el teorema de Hamilton-Cayley se sigue entonces que $A^n = 0$. El polinomio mínimo de A es $m(x) = x^k$.

8) $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

9) Por ejemplo: $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (hay más ejemplos)

10) a) El polinomio característico de ambas matrices es $p(x) = (x-2)^4$.

b) El polinomio mínimo de ambas matrices es $m(x) = (x-2)^2$.

c) Se puede ver que no existe matriz alguna P de orden 4 *inversible* tal que $AP = PB$.

11) $m(x) = x^n$

12) a) $m(x) = (x-1)^2$

b) $m(x) = (x+2)^2$

c) $m(x) = (x-1)^3$

d) $m(x) = (x-1)(x-2)^2$

e) $m(x) = (x-2)^2$

f) $m(x) = (x-1)^4$

g) $m(x) = (x-2)^4$

h) $m(x) = (x-3)^2$

15) a) $\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$

b) $\begin{bmatrix} 7 & -5 \\ -4 & 3 \end{bmatrix}$

c) $\begin{bmatrix} 3/2 & 1/2 \\ -5/2 & -1/2 \end{bmatrix}$

d) $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

e) $\begin{bmatrix} 0 & 1 & -1 \\ 2 & -2 & -1 \\ -1 & 1 & 1 \end{bmatrix}$

f) $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

SECCIÓN 4 (PÁGINA 561)

1) Siendo A simétrica, es semejante a la matriz diagonal $D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n) = diag(0, 0, \ldots, 0) = 0$, esto es, existe P inversible tal que $P^{-1}AP = 0$. De aquí que $A = POP^{-1} = 0$. Si A no es simétrica la afirmación es falsa. Por ejemplo, $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

2) Siendo A simétrica, es semejante a la matriz diagonal $D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Entonces, $0 = tr\,A = tr\,D = \lambda_1 + \lambda_2 + \ldots + \lambda_n$.

3) Siendo A y B del mismo orden, matrices simétricas, son diagonalizables, esto es, A es semejante a la matriz diagonal D en la que aparecen en su diagonal principal los valores propios de A = los valores propios de B (pues ambas tienen el mismo polinomio característico). Entonces, A y B son semejantes a la misma matriz diagonal D. Son entonces semejantes entre sí. Para ver que esta afirmación es falsa, si A y B no son simétricas vea el ejercicio 10 de la sección anterior.

4) a) $P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$, $P^tAP = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$

b) $P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$, $P^tAP = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

c) $P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$, $P^tAP = \begin{bmatrix} -4 & 0 \\ 0 & 2 \end{bmatrix}$

d) $P = \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$, $P^tAP = \begin{bmatrix} -5 & 0 \\ 0 & 5 \end{bmatrix}$

e) $P = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}$, $P^tAP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

f) $P = \begin{bmatrix} 1/3 & 2/3 & -2/3 \\ -2/3 & 2/3 & 1/3 \\ 2/3 & 1/3 & 2/3 \end{bmatrix}$, $P^tAP = \begin{bmatrix} -9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 18 \end{bmatrix}$

g) $P = \begin{bmatrix} -2/\sqrt{5} & 2/\sqrt{45} & -1/3 \\ 1/\sqrt{5} & 4/\sqrt{45} & -2/3 \\ 0 & 5/\sqrt{45} & 2/3 \end{bmatrix}$, $P^tAP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{bmatrix}$

h) $P = \begin{bmatrix} -2/3 & -1/3 & 2/3 \\ -2/3 & 2/3 & -1/3 \\ 1/3 & 2/3 & 2/3 \end{bmatrix}$, $P^tAP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{bmatrix}$

i) $P = \begin{bmatrix} 1/\sqrt{5} & 4/\sqrt{45} & 2/3 \\ 0 & 5/\sqrt{45} & -1/3 \\ -2/\sqrt{5} & 2/\sqrt{45} & 1/3 \end{bmatrix}$, $P^tAP = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 27 \end{bmatrix}$

j) $P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & -1/2 \\ 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{12} & 1/2 \\ 0 & 2/\sqrt{6} & -1/\sqrt{12} & 1/2 \\ 0 & 0 & 3/\sqrt{12} & 1/2 \end{bmatrix}$, $P^tAP = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$

5) La base es $\beta = \{(-1/\sqrt{2}, 1/\sqrt{2}, 0), (-1/\sqrt{6}, -1/\sqrt{6}, 2/\sqrt{6}), (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})\}$

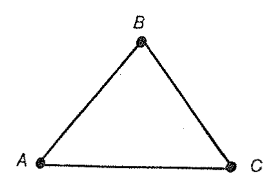
$$[T]_\beta = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

6) La base es $\beta = \{(-1/\sqrt{2}, 0, 0, 1/\sqrt{2}), (0, -1/\sqrt{2}, 1/\sqrt{2}, 0), (1/\sqrt{2}, 0, 0, 1/\sqrt{2}), (0, 1/\sqrt{2}, 1/\sqrt{2}, 0)\}$

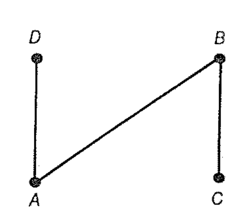
$$[T]_\beta = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

APÉNDICE (PÁGINAS 587-590)

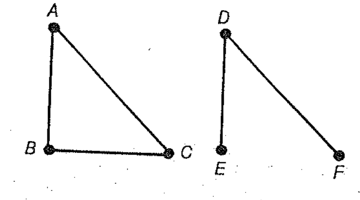
1) a)



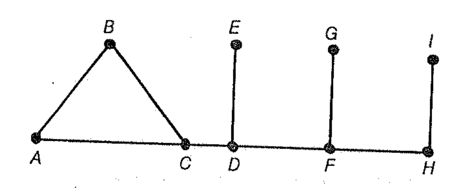
c)



b)



d)



2) Todas representan a la misma (5, 5) gráfica G en la que
$V = \{A, B, C, D, E\}$, $X = \{(A, B), (B, C), (C, D), (D, E), (E, A)\}$

3) a) La matriz de adyacencia de la primera gráfica es: $A_1 = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}$

La matriz de adyacencia de la segunda gráfica es: $A_2 = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$

b) La permutación σ es: $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 2 & 5 & 1 \end{pmatrix}$

La matriz asociada a esta permutación es: $P = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$

Se tiene $P^tA_1P = A_2$.

4) $A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$ en donde A_1 es la matriz de adyacencia de G_1 y A_2 es la matriz de adyacencia de G_2.

5) a) 2
b) 0
c) 2
d) 2
e) 0
f) 20
g) 0
h) 2
i) 61
j) 61
k) 4
l) 0

6) a) $p(\lambda) = \lambda^2(\lambda^2 - 4)$
b) $[-2, 2]$
c) $\{0 \pm 2\}$

7) a) $p(\lambda) = \lambda^2(\lambda^2 - 1)(\lambda^2 - 4)$
b) $[-3, 3]$
c) $\{0, \pm 1, \pm 2\}$

8) $p(\lambda) = \lambda^6 - 6\lambda^4 + 9\lambda^2 - 4 = (\lambda^2 - 4)(\lambda^2 - 1)^2$. Espectro $= \{\pm 1, \pm 1, \pm 2\}$. Se verifica que el espectro se encuentra en el intervalo $[-2, 2]$. Además,

$\lambda_1 + \ldots + \lambda_6 = 1 - 1 + 1 - 1 + 2 - 2 = 0$

$\lambda_1^2 + \ldots + \lambda_6^2 = 1 + 1 + 1 + 1 + 4 + 4 = 12 = 2$ (número de lados de G).

$\lambda_1^3 + \ldots + \lambda_6^3 = 1 - 1 + 1 - 1 + 8 - 8 = 0$ (no hay ciclos de longitud 3).

9) $p(\lambda) = \lambda^7 - 6\lambda^5 + 9\lambda^3 - 4\lambda = \lambda(\lambda^2 - 4)(\lambda^2 - 1)^2$

Espectro $= \{0 \pm 1, \pm 1, \pm 2\}$. Se verifica que el espectro está en $[-3, 3]$. Además,

$\lambda_1 + \ldots + \lambda_7 = 0 + 1 - 1 + 1 - 1 + 2 - 2 = 0$

$\lambda_1^2 + \ldots + \lambda_7^2 = 0 + 1 + 1 + 1 + 1 + 4 + 4 = 12 = 2$ (número de lados de G).

$\lambda_1^3 + \ldots + \lambda_7^3 = 0 + 1 - 1 + 1 - 1 + 8 - 8 = 0$ (no hay ciclos de longitud 3).

12) El polinomio característico de ambas gráficas es:
$p(\lambda) = \lambda^{10} - 10\lambda^8 + 33\lambda^6 - 44\lambda^4 + 24\lambda^2 - 4$ y su espectro es
$E = \{\pm 1, \pm 1, \pm 0.5391888, \pm 1.6751309, \pm 2.2143197\}$

CAPÍTULO SIETE

SECCIÓN 1 (PÁGINAS 602-607)

1) a) No es una forma bilineal.
b) Sí es una forma bilineal.
c) Sí es una forma bilineal.
d) No es una forma bilineal.
e) No es una forma bilineal.

5) a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

d) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

e) $\begin{bmatrix} 0 & 0 & 3 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

f) $\begin{bmatrix} 3 & 8 & -7 \\ 9 & -12 & 10 \\ 7 & 1 & -10 \end{bmatrix}$

6) a) $\begin{bmatrix} 0 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 2 & -1 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 9 & 3 \\ 0 & 18 & 6 \\ 0 & 0 & 0 \end{bmatrix}$

7) a) $\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 5 & 3 & -3 \\ 5 & 1 & -1 & 9 \\ 3 & -1 & 1 & -1 \\ -3 & 9 & -1 & 9 \end{bmatrix}$

8) a) $\begin{bmatrix} 1 & 0 & 0 & 3 & 0 & 0 \\ 0 & 1 & 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 & 0 & 3 \\ -1 & 0 & 0 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 & 2 & 0 \\ 0 & 0 & -1 & 0 & 0 & 2 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 1 & 1 & 4 & 4 & 4 \\ 1 & 2 & 2 & 5 & 8 & 8 \\ 1 & 2 & 3 & 6 & 9 & 12 \\ 0 & 1 & 2 & 7 & 10 & 13 \\ 0 & 0 & 1 & 6 & 11 & 14 \\ 0 & 0 & 0 & 5 & 10 & 15 \end{bmatrix}$

9) a) $\begin{bmatrix} 0 & 0 & 2 & 4 \\ 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 3 \\ -1 & 3 & 2 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 0 & 2 & 5 \\ 5 & 4 & 8 & 11 \\ 8 & 8 & 12 & 17 \\ 8 & 9 & 14 & 19 \end{bmatrix}$

c) $\begin{bmatrix} 19 & 17 & 9 & 9 \\ 7 & 5 & 5 & -3 \\ 7 & 9 & -3 & 5 \\ 15 & 17 & 9 & 5 \end{bmatrix}$

10) a) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ b) $\begin{bmatrix} -8/9 & 1/9 & 5/9 \\ 1/9 & -8/9 & 5/9 \\ 5/9 & 5/9 & 16/9 \end{bmatrix}$

11) Falso. Las matrices A y P^tAP no tienen, en general, la misma traza ni el mismo determinante.

12) La propiedad de simetría de A no se altera al cambiar la base de V, pues siendo A simétrica, la matriz $B = P^tAP$ también lo es ($B^t = (P^tAP)^t = (P^tA^t(P^t)^t = P^tAP = B$). Véase, por ejemplo, el ejercicio 7.

13) Las formas de los incisos c), e) y f) son no degeneradas.

14) El rango es 1.

16) El rango es 6.

17) El rango es 4.

18) El rango de la forma f (el de la matriz A) es 3, mientras que el rango de la matriz B es 2. Entonces, la matriz B no puede representar a f en alguna base de V.

19) W es no vacío, pues $0 \in W$. Tome $x, x' \in W$. Entonces $f(x + x', y) = f(x, y) + f(x', y) = 0 + 0 = 0$, lo que muestra que $x + x' \in W$. Si $c \in \mathbf{R}$ se tiene $f(cx, y) = cf(x, y) = c \cdot$

$0 = 0$ lo que muestra que $cx \in W$. Entonces, W es un subespacio de V. Sea ahora $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V, considere el operador lineal $T: V \to V$ definido como

$T(v) = \sum_{j=1}^{n} f(v, v_j)\, v_j$. Observe que $[T]_\beta = [f]^t_\beta$. Se afirma que $W = Ker\ T$.

En efecto, $v \in Ker\ T \Leftrightarrow = 0 = Tv = \sum_{j=1}^{n} f(v, v_j)\, v_j \Leftrightarrow f(v, v_j) = 0, j = 1, 2, \ldots, n \Leftrightarrow$ $f(v, u) = 0\ \forall\ u \in V \Leftrightarrow v \in W$. Al aplicar el teorema de la dimensión al operador T queda: $dim\ V = dim\ Ker\ T$ + rango de $T = dim\ W$ + rango de $[T]_\beta = dim\ W$ + rango de $[f]^t_\beta = dim\ W$ + rango de $[f]_\beta = dim\ W$ + rango de f. Entonces, rango de $f = dim\ V - dim\ W$.

20) $W = V$

21) a) $W = \{(x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_1 = 0\}$ d) $W = \{0\}$
b) $W = \{(x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_1 = x_2 = 0\}$ e) $W = \{(x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_1 = -2x_2, x_3 = 0\}$
c) $W = \{0\}$

22) $W = (\mathcal{L}(u_1))^\perp$

23) $(f_1 + 3f_2)(x, y) = 19x_1y_1 + 21x_1y_3 + 28x_2y_1 + 3x_2y_2 + 18x_2y_3 - 3x_3y_1 + 3x_3y_2 - 2x_3y_3$

$$[f_1 + 3f_2]_\beta = \begin{bmatrix} 19 & 0 & 21 \\ 28 & 3 & 18 \\ -3 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & -2 \end{bmatrix} + 3 \begin{bmatrix} 6 & 0 & 7 \\ 8 & 0 & 6 \\ -1 & 1 & 0 \end{bmatrix} = [f_1]_\beta + 3[f_2]_\beta$$

SECCIÓN 2 (PÁGINAS 629-635)

1) $dim\ S = n(n + 1)/2$

2) a) $f(x, y) = x_1y_1 + x_2y_2$
b) $f(x, y) = 2x_1y_1$
c) $f(x, y) = 6x_2y_2$
d) $f(x, y) = -\frac{5}{2}x_1y_1 - \frac{5}{2}x_2y_1$
e) $f(x, y) = 3x_1y_1 + 5x_1y_2 + 5x_2y_1 + x_2y_2$

3) a) $q(x) = [x]^t \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} [x]$

b) $q(x) = [x]^t \begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} [x]$

c) $q(x) = [x]^t \begin{bmatrix} -1 & 5/2 & 3 \\ 5/2 & 10 & -5 \\ 3 & -5 & 7 \end{bmatrix} [x]$

4) a) $q(x) = [x]^t \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} [x]$

b) $q(x) = [x]^t \begin{bmatrix} 17 & 12 & 5 \\ 12 & 8 & 4 \\ 8 & 4 & 1 \end{bmatrix} [x]$

c) $q(x) = [x]^t \begin{bmatrix} 17 & 12 & 9/2 \\ 12 & 14 & 3/2 \\ 9/2 & 3/2 & -1 \end{bmatrix} [x]$

5) $q(x) = (x \mid x)$. La matriz A es $A = ((v_i \mid v_j))_{i,j=1,\ldots,n}$ en donde $\beta = \{v_1, v_2, \ldots v_n\}$ es una base de V.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 10 \end{bmatrix} \qquad A = \begin{bmatrix} 3 & 3 & 5 \\ 3 & 5 & 9 \\ 5 & 9 & 17 \end{bmatrix}$$

7) a) 2, b) 1, c) 1, d) 2, e) 2

8) a) 3, b) 3, c) 3

11) a) $A = \begin{bmatrix} a_1^2 & a_1a_2 & a_1a_3 & \ldots & a_1a_n \\ a_1a_2 & a_2^2 & a_2a_3 & \ldots & a_2a_n \\ & & \cdots & & \\ a_1a_n & a_2a_n & a_3a_n & \ldots & a_n^2 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ & & \cdots & & \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}$

Respecto de esta base la forma cuadrática q se ve como $q(x) = (x_1')^2$ en donde $(x)_\beta = (x_1', x_2', \ldots, x_n')$. Con la perspectiva del cambio de variables, escriba q en términos de esta base equivale a hacer el cambio $x_1' = a_1x_1 + a_2x_2 + \ldots + a_nx_n$.

12) La matriz A es: $A = \begin{bmatrix} 1/n & -1/n(n-1) & -1/n(n-1) & \ldots & -1/n(n-1) \\ -1/n(n-1) & 1/n & -1/n(n-1) & \ldots & -1/n(n-1) \\ & & \cdots & & \\ -1/n(n-1) & -1/n(n-1) & -1/n(n-1) & \ldots & 1/n \end{bmatrix}$

rango de $q = n - 1$.

13) Respecto de la base (ortonormal)

$$\beta_1 = \left\{ \frac{1}{\sqrt{20 - 6\sqrt{10}}}(3 - \sqrt{10}, 1), \frac{1}{\sqrt{20 + 6\sqrt{10}}}(3 + \sqrt{10}, 1) \right\}$$

la forma q se ve como $q(x) = \frac{1 - \sqrt{10}}{2}(x_1')^2 + \frac{1 + \sqrt{10}}{2}(x_2')^2$ en donde $(x)_{\beta_1} = (x_1', x_2')$. Las ecuaciones de cambio de variables son:

$$x_1' = \frac{1}{\sqrt{20 - 6\sqrt{10}}}((3 - \sqrt{10})x_1 + x_2), \quad x_2' = \frac{1}{\sqrt{20 + 6\sqrt{10}}}((3 + \sqrt{10})x_1 + x_2)$$

Respecto de la base $\beta_2 = \left\{ (1, 0), \left(-\frac{1}{4}, 1 \right) \right\}$ la forma q se ve como $q(x) = 2(x_1'')^2 - \frac{9}{8}(x_2'')^2$ en donde $(x)_{\beta_2} = (x_1'', x_2'')$. Las ecuaciones de cambio de variables son: $x_1'' = x_1 + \frac{1}{4}x_2, \; x_2'' = x_2$.

14) Respecto de la base (ortonormal)

$$\beta_1 = \left\{ \frac{1}{\sqrt{178 + 10\sqrt{89}}}(-8, 5 + \sqrt{89}), \frac{1}{\sqrt{178 - 10\sqrt{89}}}(-8, 5 - \sqrt{89}) \right\}$$

la forma q se ve como $q(x) = \left(-\frac{3 + \sqrt{89}}{2} \right)(x_1')^2 + \left(\frac{-3 + \sqrt{89}}{2} \right)(x_2')^2$ en donde $(x)_{\beta_1} = (x_1', x_2')$. Las ecuaciones de cambio de variables son:

$$x_1' = \frac{1}{\sqrt{178 + 10\sqrt{89}}}(-8x_1 + (5 + \sqrt{89})x_2),$$

$$x_2' = \frac{1}{\sqrt{178 - 10\sqrt{89}}}(-8x_1 + (5 - \sqrt{89})x_2).$$

Respecto de la base $\beta_2 = \{(1, 0), (-4, 1)\}$ la forma q se ve como $q(x) = (x_1'')^2 - 20(x_2'')^2$ en donde $(x)_{\beta_2} = (x_1'', x_2'')$. Las ecuaciones de cambio de variables son: $x_1'' = x_1 + 4x_2, \; x_2'' = x_2$.

15) Respecto de la base (ortonormal)

$$\beta_1 = \left\{ \frac{1}{\sqrt{2}}(-1, 0, 1), \frac{1}{\sqrt{6}}(1, -2, 1), \frac{1}{\sqrt{3}}(1, 1, 1) \right\}$$

La forma q se ve como $q(x) = 6(x_1')^2 + 6(x_2')^2 + 9(x_3')^2$ en donde $(x)_{\beta_1} = (x_1', x_2', x_3')$. Las ecuaciones de cambio de variables son:

$$x_1' = \frac{1}{\sqrt{2}}(-x_1 + x_3), x_2' = \frac{1}{\sqrt{6}}(x_1 - 2x_2 + x_3), x_3' = \frac{1}{\sqrt{3}}(x_1 + x_2 + x_3).$$

Respecto de la base $\beta_2 = \left\{ (1, 0, 0), \left(-\frac{1}{7}, 1, 0 \right), \left(-\frac{1}{8}, -\frac{1}{8}, 1 \right) \right\}$ la forma q se ve como $q(x) = 7(x_1'')^2 + \frac{48}{7}(x_2'')^2 + \frac{27}{4}(x_3'')^2$ en donde $(x)_{\beta_2} = (x_1'', x_2'', x_3'')$. Las ecuaciones de cambio de variables son: $x_1'' = x_1 + \frac{1}{7}x_2 + \frac{1}{7}x_3, x_2'' = x_2 + \frac{1}{8}x_3, x_3'' = x_3$.

16) Respecto de la base (ortonormal)

$$\beta_1 = \left\{ \frac{1}{3}(-2, 2, 1), \frac{1}{3}(2, 1, 2), \frac{1}{3}(-1, -2, 2) \right\}$$

la forma q se ve como $q(x) = (x_1')^2 + 4(x_2')^2 + 7(x_3')^2$ en donde $(x)_{\beta_1} = (x_1', x_2', x_3')$. Las ecuaciones de cambio de variables son:

$$x_1' = \frac{1}{3}(-2x_1 + 2x_2 + x_3), x_2' = \frac{1}{3}(2x_1 + x_2 + 2x_3), x_3' = \frac{1}{3}(-x_1 - 2x_2 + 2x_3)$$

Respecto de la base $\beta_2 = \left\{ (1, 0, 0), \left(-\frac{2}{3}, 1, 0 \right), \left(-\frac{1}{2}, \frac{3}{4}, 1 \right) \right\}$ la forma q se ve como $q(x) = 3(x_1'')^2 + \frac{8}{3}(x_2'')^2 + \frac{7}{2}(x_3'')^2$ en donde $(x)_{\beta_2} = (x_1'', x_2'', x_3'')$. Las ecuaciones de cambio de variables son: $x_1'' = x_1 + \frac{2}{3}x_2, x_2'' = x_2 + \frac{3}{4}x_3, x_3'' = x_3$.

17) Respecto de la base (ortonormal)

$$\beta_1 = \left\{\frac{1}{3}(2,1,2), \frac{1}{\sqrt{5}}(1,-2,0), \frac{1}{\sqrt{45}}(-4,-2,5)\right\}$$

la forma q se ve como $q(x) = -2(x_1')^2 + 7(x_2')^2 + 7(x_3')^2$ en donde $(x)_{\beta_1} = (x_1', x_2', x_3')$.
Las ecuaciones de cambio de variables son:

$$x_1' = \frac{1}{3}(2x_1 + x_2 + 2x_3),\ x_2' = \frac{1}{\sqrt{5}}(x_1 - 2x_2),\ x_3' = \frac{1}{\sqrt{45}}(-4x_1 - 2x_2 - 5x_3)$$

Respecto de la base $\beta_2 = \left\{(1,0,0), \left(\frac{2}{3}, 1, 0\right), (2,1,1)\right\}$ la forma q se ve como $q(x) = 3(x_1'')^2 + \frac{14}{3}(x_2'')^2 - 7(x_3'')^2$ en donde $(x)_{\beta_2} = (x_1'', x_2'', x_3'')$. Las ecuaciones de cambio de variables son: $x_1'' = x_1 - \frac{2}{3}x_2 - \frac{4}{3}x_3,\ x_2'' = x_2 - x_3,\ x_3'' = x_3$.

18) Respecto de la base (ortonormal)

$$\beta_1 = \left\{\frac{1}{3}(2,2,-1), \frac{1}{3}(-1,2,2), \frac{1}{3}(2,-1,2)\right\}$$

la forma q se ve como $q(x) = 9(x_1')^2 + 18(x_2')^2 - 9(x_3')^2$ en donde $(x)_{\beta_1} = (x_1', x_2', x_3')$.
Las ecuaciones de cambio de variables son:

$$x_1' = \frac{1}{3}(2x_1 + 2x_2 - x_3),\ x_2' = \frac{1}{3}(-x_1 + 2x_2 + 2x_3),\ x_3' = \frac{1}{3}(2x_1 - x_2 + 2x_3)$$

Respecto de la base $\beta_2 = \left\{(1,0,0), \left(\frac{-8}{11}, 1, 0\right), (10,-14,1)\right\}$ la forma q se ve como $q(x) = 11(x_1'')^2 - \frac{9}{11}(x_2'')^2 + 162(x_3'')^2$ en donde $(x)_{\beta_2} = (x_1'', x_2'', x_3'')$. Las ecuaciones de cambio de variables son: $x_1'' = x_1 + \frac{8}{11}x_2 + \frac{2}{11}x_3,\ x_2'' = x_2 + 14x_3,\ x_3'' = x_3$.

19) Respecto de la base (ortonormal)

$$\beta_1 = \left\{\frac{1}{3}(-1,2,2), \frac{1}{\sqrt{5}}(2,1,0), \frac{1}{\sqrt{45}}(2,-4,5)\right\}$$

la forma q se ve como $q(x) = -2(x_1')^2 + 7(x_2')^2 + 7(x_3')^2$ en donde $(x)_{\beta_1} = (x_1', x_2', x_3')$.
Las ecuaciones de cambio de variables son:

$$x_1' = \frac{1}{3}(-x_1 + 2x_2 + 2x_3),\ x_2' = \frac{1}{\sqrt{5}}(2x_1 + x_2),\ x_3' = \frac{1}{\sqrt{45}}(2x_1 - 4x_2 + 5x_3)$$

Respecto de la base $\beta_2 = \left\{(1,0,0), \left(-\frac{1}{3}, 1, 0\right), (-1,2,1)\right\}$ la forma q se ve como $q(x) = 6(x_1'')^2 + \frac{7}{3}(x_2'')^2 - 7(x_3'')^2$ en donde $(x)_{\beta_2} = (x_1'', x_2'', x_3'')$. Las ecuaciones de cambio de variables son: $x_1'' = x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_3,\ x_2'' = x_2 - 2x_3,\ x_3'' = x_3$.

20) Respecto de la base (ortonormal)

$$\beta_1 = \left\{\frac{1}{3}(-1,-2,2), \frac{1}{3}(-2,2,1), \frac{1}{3}(2,1,2)\right\}$$

la forma q se ve como $q(x) = -(x_1')^2 + 11(x_2')^2 + 5(x_3')^2$ en donde $(x)_{\beta_1} = (x_1', x_2', x_3')$.
Las ecuaciones de cambio de variables son:

$$x_1' = \frac{1}{3}(-x_1 - 2x_2 + 2x_3),\ x_2' = \frac{1}{3}(-2x_1 + 2x_2 + x_3),\ x_3' = \frac{1}{3}(2x_1 + x_2 + 2x_3)$$

Respecto de la base $\beta_2 = \left\{(1,0,0), \left(\frac{4}{7}, 1, 0\right), \left(-\frac{16}{19}, -\frac{28}{19}, 1\right)\right\}$ la forma q se ve como $q(x) = 7(x_1'')^2 + \frac{19}{7} - (x_2'')^2 - \frac{55}{9}(x_3'')^2$ en donde $(x)_{\beta_2} = (x_1'', x_2'', x_3'')$. Las ecuaciones de cambio de variables son: $x_1'' = x_1 - \frac{4}{7}x_2,\ x_2'' = x_2 + \frac{28}{19}x_3,\ x_3'' = x_3$.

21) Respecto de la base (ortonormal)

$$\beta_1 = \left\{\frac{1}{2}(1,1,1,1), \frac{1}{2}(1,-1,-1,1), \frac{1}{2}(1,1,-1,-1), \frac{1}{2}(1,-1,1,-1)\right\}$$

la forma q se ve como $q(x) = 4(x_2')^2 + 12(x_2')^2 - 4(x_3')^2 + 8(x_4')^2$ en donde $(x)_{\beta_1} = (x_1', x_2', x_3', x_4')$. Las ecuaciones de cambio de variables son:

$$x_1 = \frac{1}{2}(x_1 + x_2 + x_3 + x_4) \qquad x_3 = \frac{1}{2}(x_1 + x_2 - x_3 - x_4)$$

$$x_2 = \frac{1}{2}(x_1 - x_2 - x_3 + x_4) \qquad x_4 = \frac{1}{2}(x_1 - x_2 + x_3 - x_4)$$

Respecto de la base $\beta_2 = \left\{(1,0,0), \left(\frac{4}{5}, 1, 1, 0\right), \left(-\frac{3}{4}, -\frac{11}{16}, \frac{5}{16}, 0\right), \left(\frac{11}{5}, \frac{13}{5}, -1, 1\right)\right\}$ la forma q se ve como $q(x) = 5(x_1'')^2 + \frac{64}{5}(x_2'')^2 - \frac{5}{4}(x_3'')^2 + \frac{96}{5}(x_4'')^2$ en donde $(x)_{\beta_2} = (x_1'', x_2'', x_3'', x_4'')$. Las ecuaciones de cambio de variables son: $x_1'' = x_1 - x_2 + \frac{1}{5}x_3 + \frac{3}{5}x_4,\ x_2'' = \frac{5}{16}x_2 + \frac{11}{16}x_3 - \frac{1}{8}x_4,\ x_3'' = -x_2 + x_3 + \frac{18}{5}x_4,\ x_4'' = x_4$.

22) Respecto de la base (ortonormal)

$$\beta_1 = \left\{\frac{1}{2}(1,-1,-1,1), \frac{1}{\sqrt{2}}(1,0,1,0), \frac{1}{\sqrt{2}}(0,1,0,1), \frac{1}{2}(-1,-1,1,1)\right\}$$

la forma q se ve como $q(x) = -(x_1')^2 + (x_2')^2 + (x_3')^2 + 3(x_4')^2$ en donde $(x)_{\beta_1} = (x_1', x_2', x_3', x_4')$. Las ecuaciones de cambio de variables son:

$$x_1' = \frac{1}{2}(x_1 - x_2 - x_3 + x_4) \qquad x_3 = \frac{1}{\sqrt{2}}(x_2 + x_4)$$

$$x_2' = \frac{1}{2}(x_1 + x_3) \qquad x_4 = \frac{1}{2}(-x_1 - x_2 + x_3 + x_4)$$

} la for-

$, x''_3, x''_4)$.

Las ecuaciones de cambio de variables son: $x''_1 = x_1 + x_2 - x_4$, $x''_2 = x_2 - 2x_4$, $x''_3 = x_3 - x_2 + x_4$, $x''_4 = x_4$.

23) Respecto de la base (ortonormal)

$$\beta_1 = \left\{\frac{1}{2}(1,-1,-1,1), \frac{1}{2}(-1,-1,1,1), \frac{1}{2}(1,1,1,1), \frac{1}{2}(-1,1,-1,1)\right\}$$

la forma q se ve como $q(x) = -3(x'_1)^2 - (x'_2)^2 + (x'_3)^2 + 7(x'_4)^2$ en donde $(x)_{\beta_1} = (x'_1, x'_2, x'_3, x'_4)$. Las ecuaciones de cambio de variables son: $x'_1 = \frac{1}{2}(x_1 - x_2 - x_3 + x_4)$, $x'_2 = \frac{1}{2}(-x_1 - x_2 + x_3 + x_4)$, $x'_3 = \frac{1}{2}(x_1 + x_2 + x_3 + x_4)$, $= x'_4 = \frac{1}{2}(-x_1 + x_2 - x_3 + x_4)$.

Respecto de la base $\beta_2 = \left\{(1,0,0,0), (-2,1,1,0), \left(-\frac{2}{3}, -\frac{7}{6}, -\frac{1}{6}, 0\right), (-8,-13,-1,1)\right\}$ la forma q se ve como $q(x) = (x''_1)^2 - 6(x''_2)^2 + \frac{1}{6}(x''_3)^2 - 21(x''_4)^2$ en donde $(x)_{\beta_2} = (x''_1, x''_2, x''_3, x''_4)$. Las ecuaciones de cambio de variables son: $x''_1 = x_1 - x_2 + 3x_3 - 2x_4$, $x''_2 = -\frac{1}{6}x_2 + \frac{7}{6}x_3 - x_4$, $x''_3 = -x_2 + x_3 - 12x_4$, $x''_4 = x_4$.

24) Respecto de la base (ortonormal)

$$\beta_1 = \left\{\frac{1}{2}(-1,-1,1,1), \frac{1}{2}(1,-1,-1,1), \frac{1}{2}(1,-1,1,-1), \frac{1}{2}(1,1,1,1)\right\}$$

la forma q se ve como $q(x) = -5(x'_1)^2 - 3(x'_2)^2 + 3(x'_3)^2 + 5(x'_4)^2$ en donde $(x)_{\beta_1} = (x'_1, x'_2, x'_3, x'_4)$. Las ecuaciones de cambio de variables son: $x'_1 = \frac{1}{2}(-x_1 - x_2 + x_3 + x_4)$, $x'_2 = \frac{1}{2}(x_1 - x_2 - x_3 + x_4)$, $x'_4 = \frac{1}{2}(x_1 - x_2 + x_3 - x_4)$, $x'_4 = \frac{1}{2}(x_1 + x_2 + x_3 + x_4)$. Respecto de la base

$$\beta_2 = \left\{(1,0,1,0), \left(-\frac{1}{8}, 1, -\frac{1}{8}, 0\right), \left(-\frac{1}{2}, \frac{1}{8}, \frac{1}{2}, -\frac{1}{8}\right), (-4,31,-4,1)\right\}$$

la forma q se ve como $q(x) = 8(x''_1)^2 - \frac{1}{8}(x''_2)^2 - \frac{15}{8}(x''_3)^2 + 120(x''_4)^2$ en donde $(x)_{\beta_2} = (x''_1, x''_2, x''_3, x''_4)$. Las ecuaciones de cambio de variables son: $x''_1 = \frac{1}{2}x_1 + \frac{1}{8}x_2 + \frac{1}{2}x_3 + \frac{1}{8}x_4$, $x''_2 = 4x_1 + x_2 - 4x_3 - 31x_4$, $x''_3 = x_3 - x_1$, $x''_4 = -\frac{1}{8}x_1 + \frac{1}{8}x_3 + x_4$.

25) Respecto de la base (ortonormal)

$$\beta_1 = \left\{\frac{1}{\sqrt{13}}(0,0,0,-3,2), \frac{1}{\sqrt{5}}(1,2,0,0,0), \frac{1}{\sqrt{5}}(-2,1,0,0,0), (0,0,1,0,0), \frac{1}{\sqrt{13}}(0,0,0,2,3)\right\}$$